MANUAL BÁSICO SPSS

Elaborado por: José Aurelio Pina Romero
Departamento de Enfermería Comunitaria
Universidad de Alicante
Enero 2020

ÍNDICE

1. INTRODUCCIÓN A SPSS **3**
1.1 VISOR DE DATOS 3
1.2 VISOR DE VARIABLES 4
1.3 VISOR DE RESULTADOS 12
1.4 GUARDAR UNA BASE DE DATOS 14
1.5 ABRIR UNA BASE DE DATOS 15

2. TRANSFORMACIÓN DE LOS DATOS **16**
2.1 CALCULO DE UNA VARIABLE NUEVA 16
2.2 RECODIFICACIÓN EN DISTINTAS VARIABLES 16
2.2 RECODIFICACIÓN EN DISTINTAS VARIABLES. 18
2.3 PONDERAR CASOS 21
2.4 DIVIDIR/SEGMENTAR ARCHIVO 23

3. ANÁLISIS ESTADÍSTICO **25**
3.1 MEDIDAS DE TENDENCIA CENTRAL 25
3.2 MEDIDAS DE DISPERSIÓN 27
3.3 MEDIDAS DE FORMA 29
3.4 MEDIDAS DE POSICIÓN 31
3.5 RESULTADOS 33

4. GRÁFICOS **34**
4.1 GRÁFICOS DE SECTORES. 34
4.2 GRÁFICOS DE BARRAS. 36
4.4 HISTOGRAMA. 37
4.5 GRÁFICOS DE DISPERSIÓN. 39

5. ANÁLISIS DE VARIABLES CATEGÓRICAS (TABLAS DE CONTINGENCIA. **43**

6. ANÁLISIS DE CORRELACIÓN **46**

7. ANÁLISIS DE REGRESIÓN LINEAL **49**

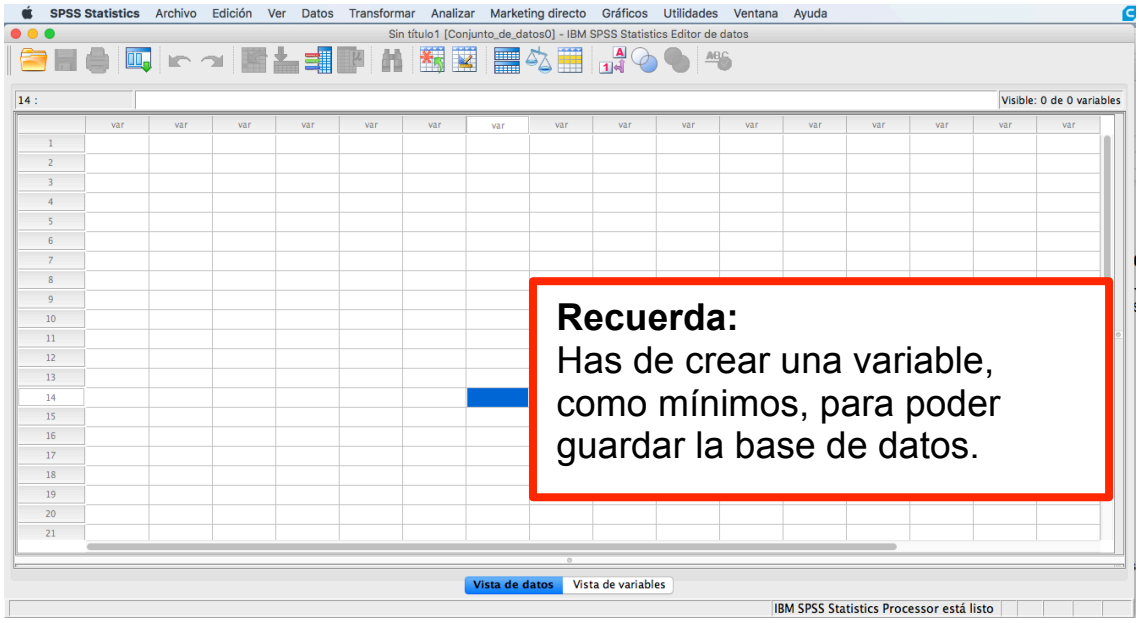
8. INTERVALOS DE CONFIANZA **52**
8.1 INTERVALO DE CONFIANZA PARA UNA MEDIA 52
8.2 INTERVALO DE CONFIANZA PARA LA DIFERENCIA DE MEDIAS 54
8.3 INTERVALO DE CONFIANZA PARA LA DIFERENCIA DE MEDIAS EN MUESTRAS RELACIONADAS 56

9. CONTRASTE DE HIPÓTESIS **57**
9.1 CONTRASTE DE HIPÓTESIS PARA UNA MEDIA 57
9.2 CONTRASTE DE HIPÓTESIS PARA UNA DIFERENCIA DE MEDIAS 59
9.3 CONTRASTE DE HIPÓTESIS CHI-CUADRADO 61

1. Introducción a SPSS

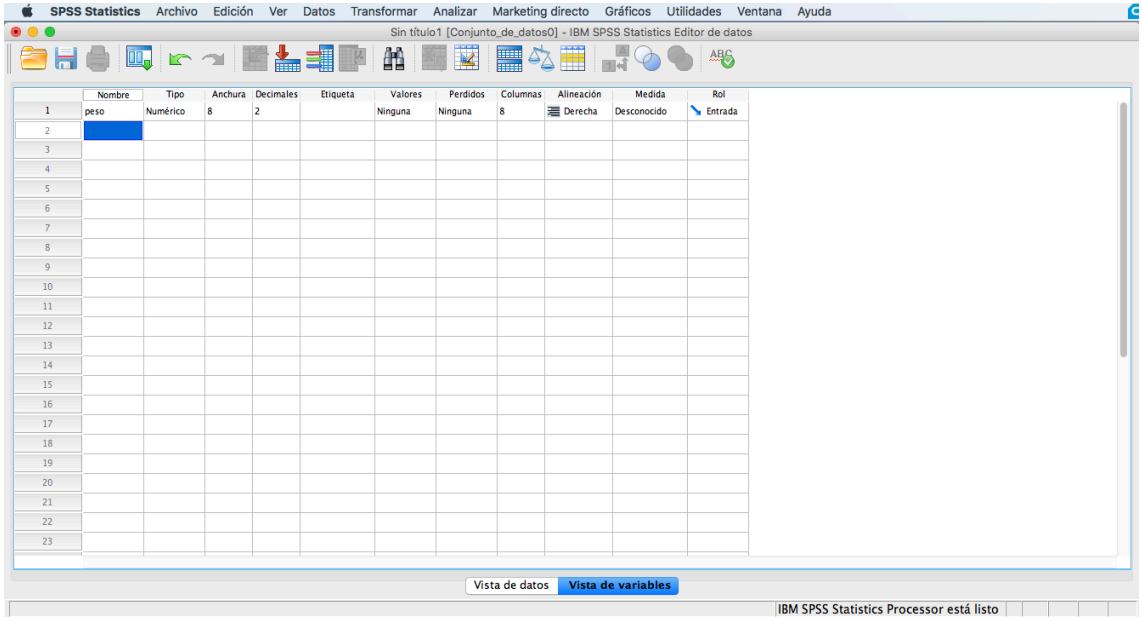
SPSS es uno de los programas que más se utilizan en el analizar datos. Una de las características fundamentales de SPSS es su facilidad de uso, junto a la potencia e integridad del software.

1.1 Visor de datos



En la que introduciremos los datos de las variables a trabajar. Similar a una hoja de cálculo, las columnas representan variables y las filas los datos.

1.2 Visor de variables



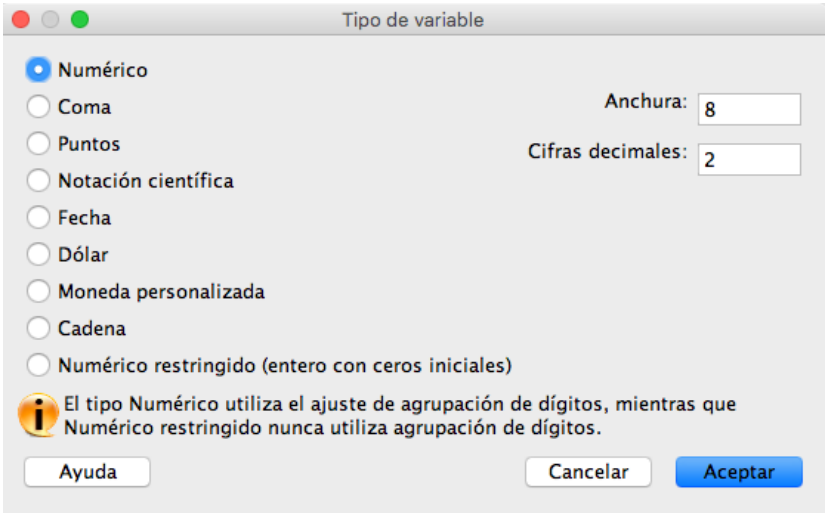
En el que definiremos las variables que componen la encuesta del estudio, es decir sus atributos.

	Nombre	Tipo	Anchura	Decimales	Etiqueta	Valores	Perdidos	Columnas	Alineación	Medida	Rol
1	peso	Numérico	8	2		Ninguna	Ninguna	8	Derecha	Desconocido	Entrada
2											

Parámetros de la variable

Nombre: Nombre de la variable.

Tipo:



Numérico. Una variable cuyos valores son números. Los valores se muestran en formato numérico estándar. El Editor de datos acepta valores numéricos en formato estándar o en notación científica.

Coma. Una variable numérica cuyos valores se muestran con comas que delimitan cada tres posiciones y con el punto como delimitador decimal. El Editor de datos acepta valores numéricos para este tipo de variables con o sin comas, o bien en notación científica. Los valores no pueden contener comas a la derecha del indicador decimal.

Punto. Una variable numérica cuyos valores se muestran con puntos que delimitan cada tres posiciones y con la coma como delimitador decimal. El Editor de datos acepta valores numéricos para este tipo de variables con o sin puntos, o bien en notación científica. Los valores no pueden contener puntos a la derecha del indicador decimal.

Notación científica. Una variable numérica cuyos valores se muestran con una E intercalada y un exponente con signo que representa una potencia de base 10. El Editor de datos acepta para estas variables valores numéricos con o sin el exponente. El exponente puede aparecer precedido por una E o una D con un signo opcional, o bien sólo por el signo (por ejemplo, 123, 1,23E2, 1,23D2, 1,23E+2 y 1,23+2).

Fecha. Una variable numérica cuyos valores se muestran en uno de los diferentes formatos de fecha-calendario u hora-reloj. Seleccione un formato de la lista. Puede introducir las fechas utilizando como delimitadores: barras, guiones, puntos, comas o espacios. El rango de siglo para los valores de año de dos dígitos está determinado por la configuración de las opciones (en el menú Edición, seleccione **Opciones** y, a continuación, pulse en la pestaña **Datos**).

Dólar. Una variable numérica que se muestra con un signo dólar inicial ($), comas que delimitan cada tres posiciones y un punto como delimitador decimal. Se pueden introducir valores de datos con o sin el signo dólar inicial.

Moneda personalizada. Una variable numérica cuyos valores se muestran en uno de los formatos de moneda personalizados que se hayan definido previamente en la pestaña Moneda del cuadro de diálogo Opciones. Los caracteres definidos en la moneda personalizada no se pueden emplear en la introducción de datos pero sí se mostrarán en el Editor de datos.

Cadena. Una variable cuyos valores no son numéricos y, por lo tanto, no se utilizan en los cálculos. Los valores pueden contener cualquier carácter siempre que no se exceda la longitud definida. Las mayúsculas y las minúsculas se consideran diferentes. Este tipo también se conoce como variable alfanumérica.

Numérico restringido. Una variable cutos valores están restringidos para enteros no negativos. Los valores aparecen con los ceros iniciales llenando el ancho máximo de la variable. Los valores se pueden introducir en notación científica.

Anchura - Decimales:

Se puede especificar un número de caracteres para el ancho de la columna. Los anchos de columna también se pueden cambiar en la Vista de datos pulsando y arrastrando los bordes de las columnas.

- El ancho de columna para fuentes proporcionales se basa en el ancho medio de los caracteres. Dependiendo de los caracteres utilizados en el valor, se mostrarán más o menos caracteres con el ancho especificado.

- El ancho de columna afecta sólo a la presentación de valores en el Editor de datos. Al cambiar el ancho de columna no se cambia el ancho definido de una variable.

Etiqueta:

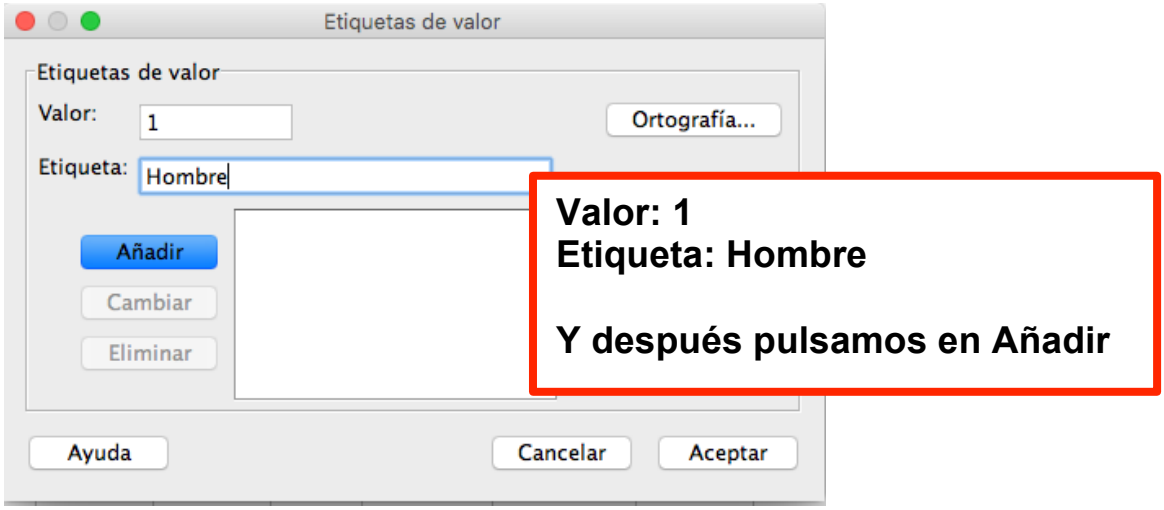
Puede asignar etiquetas de variable descriptivas de hasta 256 caracteres de longitud (128 caracteres en los idiomas de doble byte). Las etiquetas de variable pueden contener espacios y caracteres reservados que no se admiten en los nombres de variable.

Valores:

Puede asignar etiquetas de valor descriptivas a cada valor de una variable. Este proceso es especialmente útil si el archivo de datos utiliza códigos numéricos para representar categorías que no son numéricas (por ejemplo, códigos 1 y 2 para *hombre* y *mujer*).

- Las etiquetas de valor se guardan con el archivo de datos. No es necesario volver a definir las etiquetas de valor cada vez que se abre un archivo de datos.

- Las etiquetas de valor pueden ocupar hasta 120 bytes.



Perdidos:

Valores perdidos define los valores de los datos definidos como **perdidos por el usuario**. Por ejemplo, es posible que quiera distinguir los datos perdidos porque un encuestado se niegue a responder de los datos perdidos porque la pregunta no afecta a dicho encuestado. Los valores de datos que se especifican como perdidos por el usuario aparecen marcados para un tratamiento especial y se excluyen de la mayoría de los cálculos.

- Las especificaciones de valores perdidos definidos por el usuario se guardan junto con el archivo de datos. No es necesario volver a definir los valores definidos como perdidos por el usuario cada vez que se abre un archivo de datos.

- Se pueden introducir hasta tres valores perdidos (individuales) de tipo discreto, un rango de valores perdidos o un rango más un valor de tipo discreto.

- Sólo pueden especificarse rangos para las variables numéricas.

- Se considera que son válidos todos los valores de cadena, incluidos los valores vacíos o nulos, a no ser que se definan explícitamente como perdidos.

- Los valores perdidos de las variables de cadena no pueden tener más de ocho bytes. (No hay ningún límite respecto al ancho definido de la variable de cadena, pero los valores definidos como perdidos no pueden tener más de ocho bytes.)

- Para definir como perdidos los valores nulos o vacíos de una variable de cadena, escriba un espacio en blanco en uno de los campos debajo de la selección **Valores perdidos discretos**.

Columnas

Se puede especificar un número de caracteres para el ancho de la columna. Los anchos de columna también se pueden cambiar en la Vista de datos pulsando y arrastrando los bordes de las columnas.

- El ancho de columna para fuentes proporcionales se basa en el ancho medio de los caracteres. Dependiendo de los caracteres utilizados en el valor, se mostrarán más o menos caracteres con el ancho especificado.
- El ancho de columna afecta sólo a la presentación de valores en el Editor de datos. Al cambiar el ancho de columna no se cambia el ancho definido de una variable.

Alineación:

La alineación controla la presentación de los valores de los datos y/o de las etiquetas de valor en la Vista de datos. La alineación por defecto es a la derecha para las variables numéricas y a la izquierda para las variables de cadena. Este ajuste sólo afecta a la presentación en la Vista de datos.

Cambio de la alineación de los valores en la Vista de datos

- Haga que el editor de datos sea la ventana activa.
- Pulse dos veces en un nombre de variable en la parte superior de la columna en la Vista de datos o bien pulse en la pestaña **Vista de variables**.
- Seleccione la alineación (izquierda, derecha, central) de la lista en la casilla *Alineación* de la variable.

Medida:

Puede especificar el nivel de medida como Escala (datos numéricos de una escala de intervalo o de razón), Ordinal o Nominal. Los datos nominales y ordinales pueden ser de cadena (alfanuméricos) o numéricos.

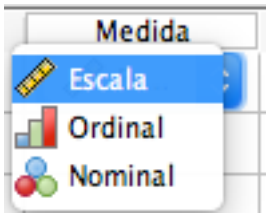
- **Nominal.** Una variable se puede tratar como nominal si sus valores representan categorías que no obedecen a una ordenación intrínseca (por ejemplo, el departamento de la empresa en el que trabaja un empleado). Algunos ejemplos de variables nominales son: región, código postal o confesión religiosa, sexo, etc...

- **Ordinal.** Una variable puede tratarse como ordinal cuando sus valores representan categorías con alguna ordenación intrínseca (por ejemplo, los niveles de satisfacción con un servicio, que vayan desde muy insatisfecho hasta muy satisfecho). Entre los ejemplos de variables ordinales se incluyen escalas de actitud que representan el grado de satisfacción o confianza y las puntuaciones de evaluación de las preferencias.

- **Escala.** Una variable puede tratarse como escala (continua) cuando sus valores representan categorías ordenadas con una métrica con significado, por lo que son adecuadas las comparaciones de distancia entre valores. Son ejemplos de variables de escala: la edad en años y los ingresos en dólares.

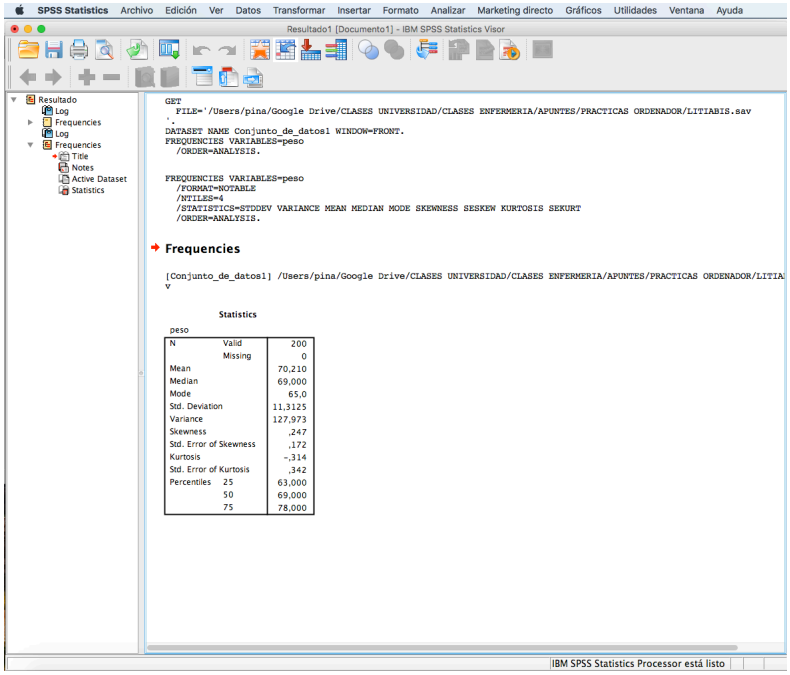
Nota: Para variables de cadena ordinales, se asume que el orden alfabético de los valores de cadena indica el orden correcto de las categorías. Por ejemplo, en una variable de cadena cuyos valores sean *bajo*, *medio*, *alto*, se interpreta el orden de las categorías como *alto*, *bajo*, *medio* (orden que no es el correcto). Por norma general, se puede indicar que es más fiable utilizar códigos numéricos para representar datos ordinales.

Para nuevas variables numéricas creadas con transformaciones, los datos de orígenes externos y los archivos de datos de IBM® SPSS® Statistics creados antes de la versión 8, el nivel de medición por defecto está determinado por las condiciones de la tabla siguiente. Las condiciones se evalúan en el orden de la tabla. Se aplicará el nivel de medición de la primera condición que coincida con los datos.

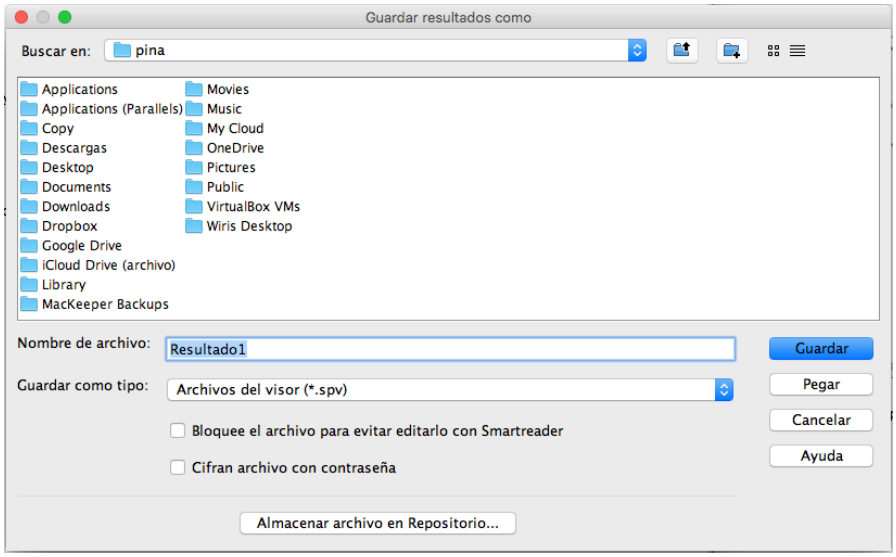


1.3 Visor de Resultados

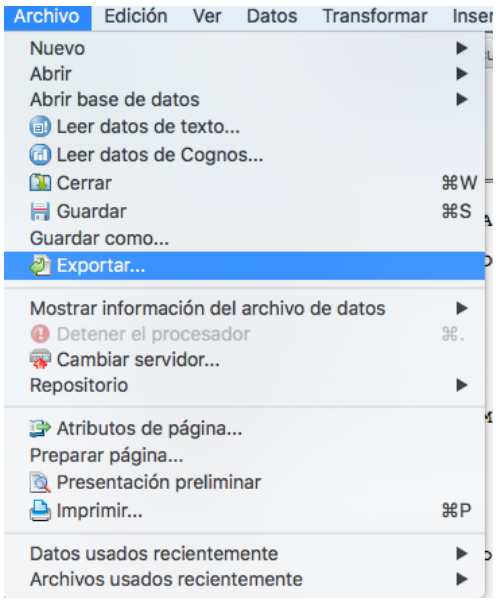
Corresponde a una ventana donde el software despliega los resultados que se obtienen de los distintos análisis realizados. En la sección izquierda, el software mostrará los títulos de los distintos análisis y los desplegará como una diagrama de árbol.

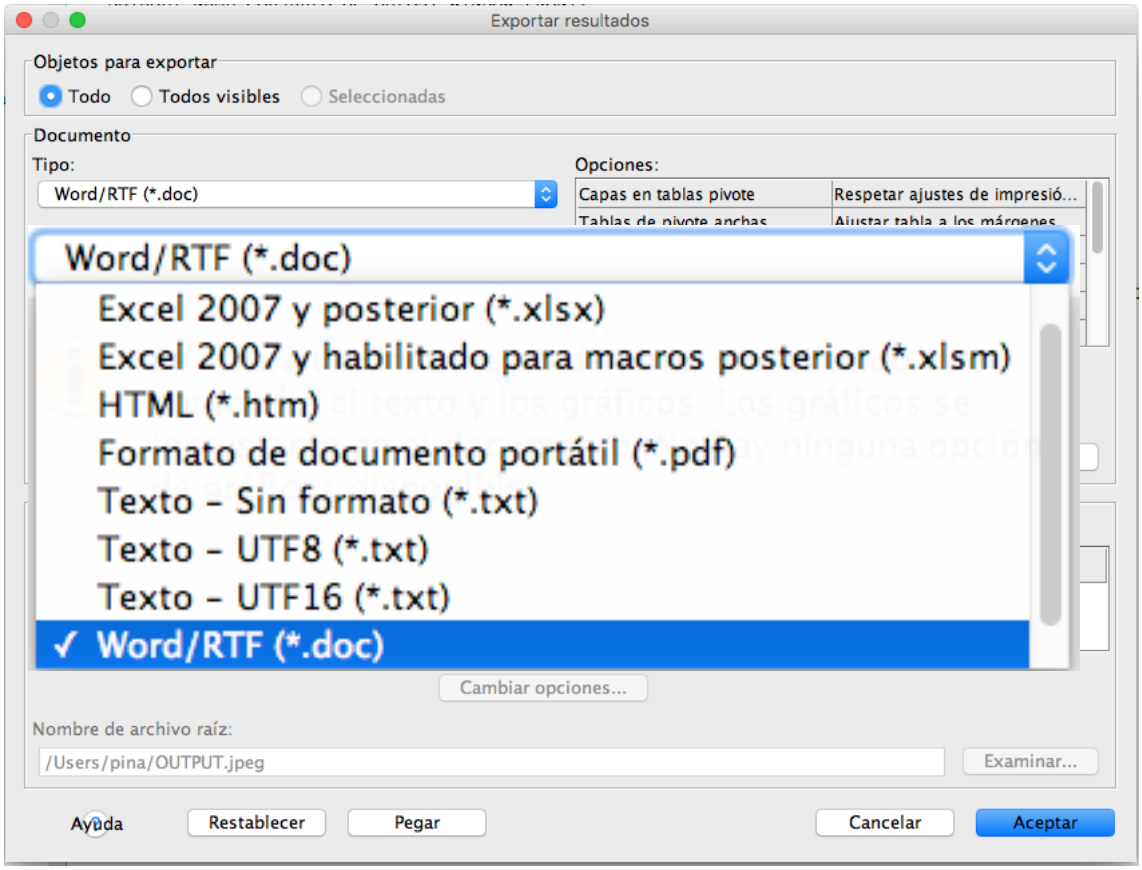


Se puede guardar en, y su formato es “.spv”

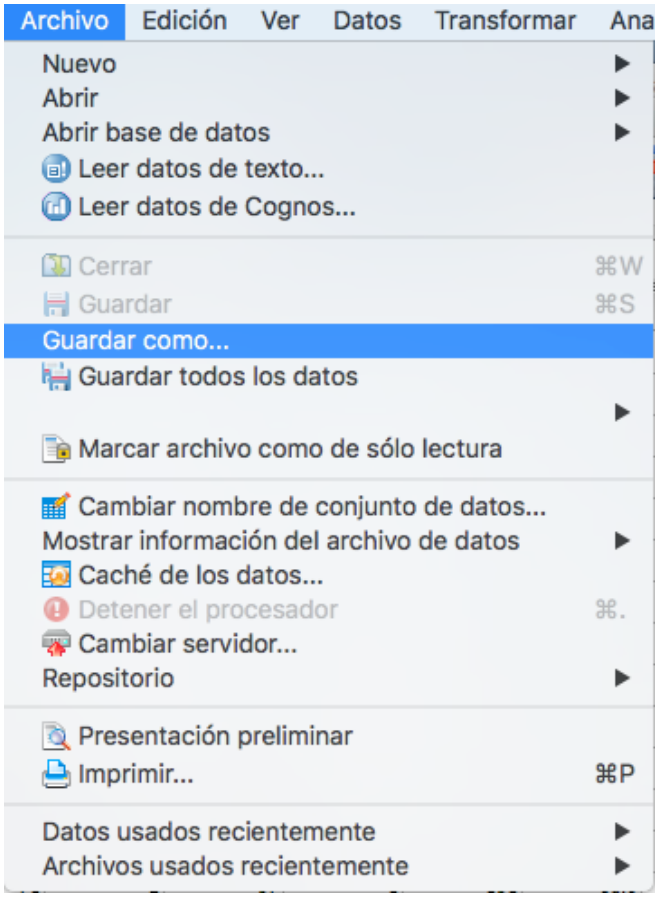


Y además, también se puede exportar.

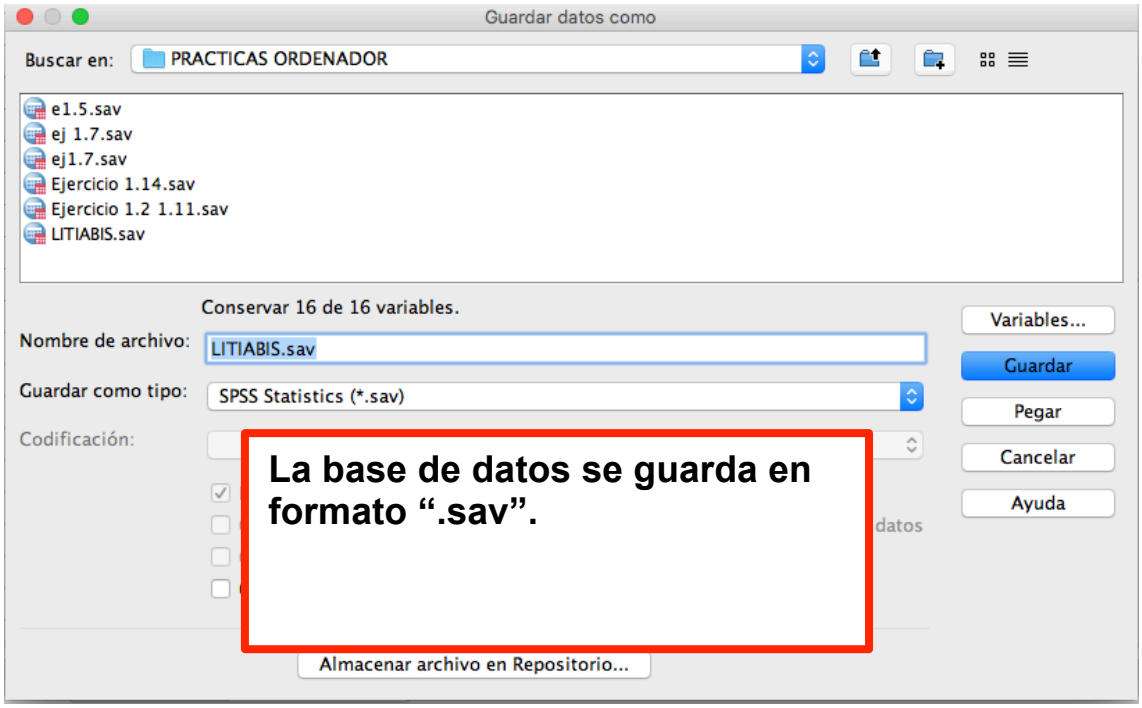




1.4 Guardar una base de datos

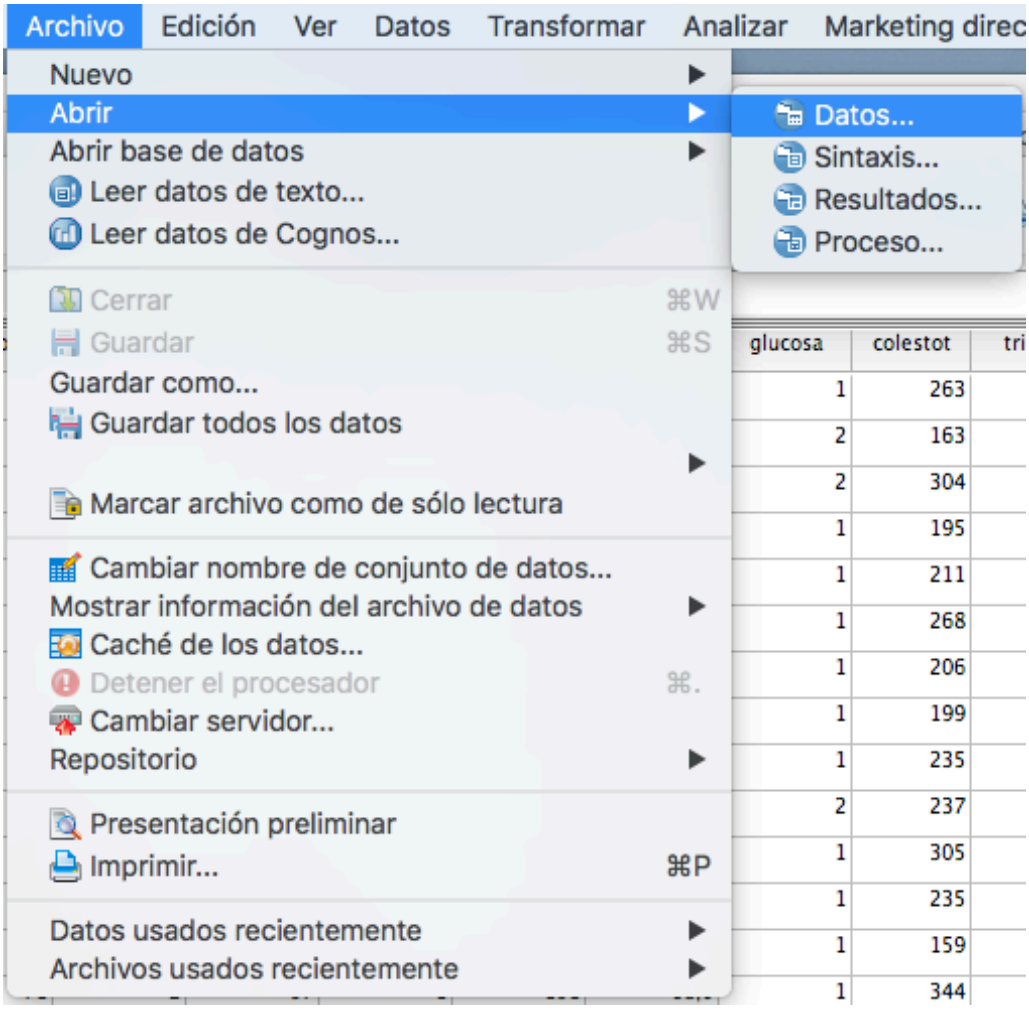


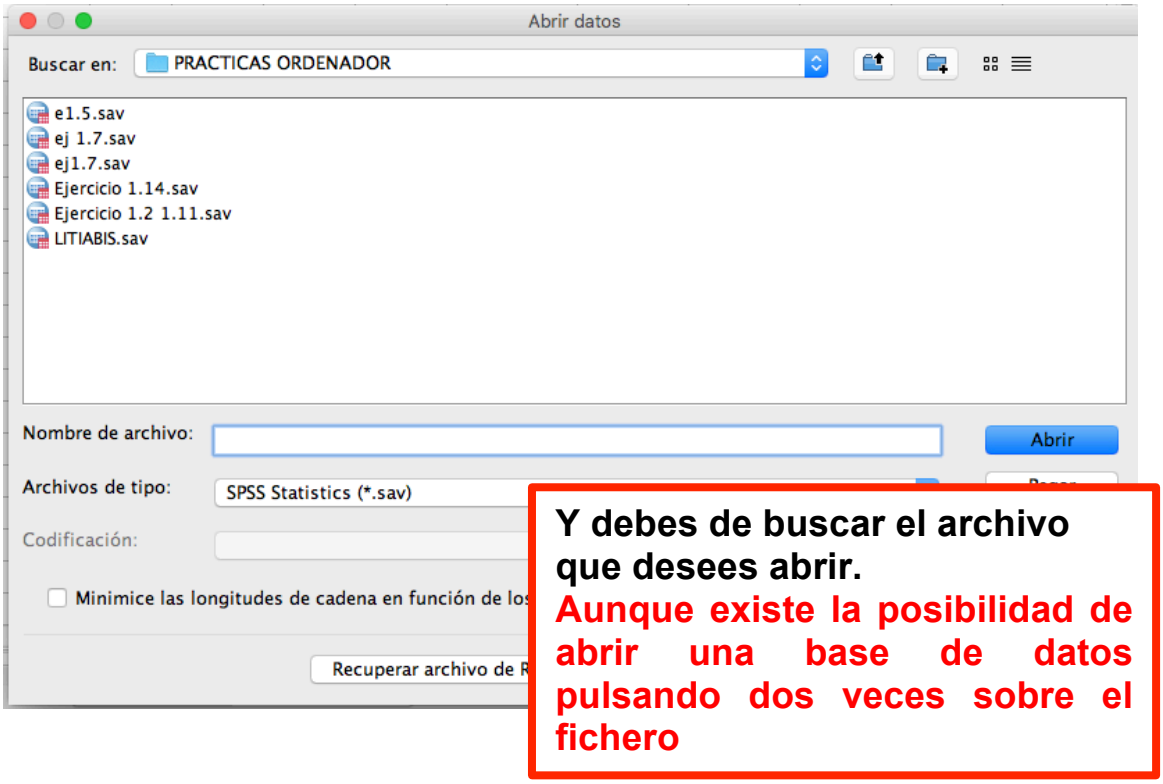
Ahora has de guardar en el directorio que consideres, ya que debes de trabajar con esa base de datos en el futuro.



La base de datos se guarda en formato ".sav".

1.5 Abrir una base de datos

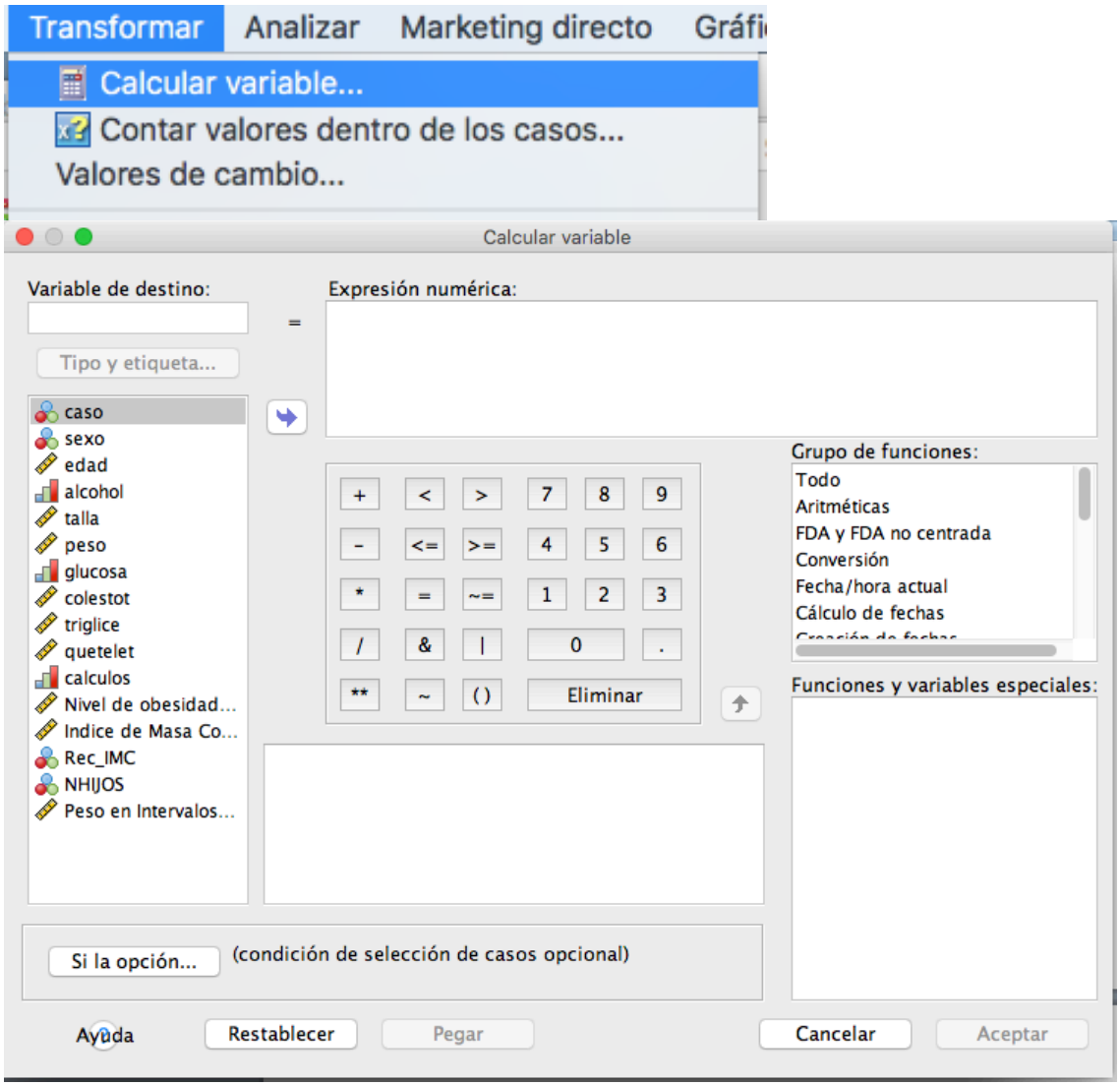




2. Transformación de los datos

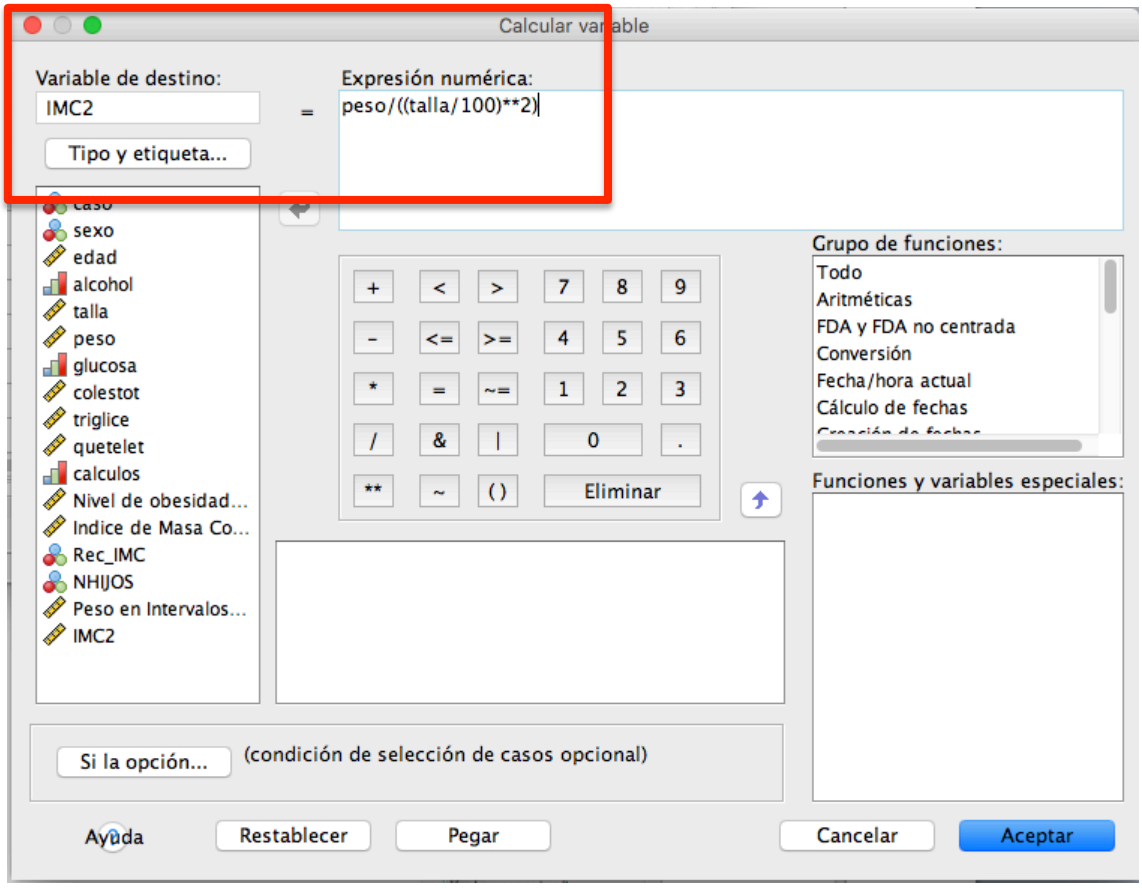
2.1 Calculo de una variable nueva

A partir de una o varias variables, se pueden crear otras variables. Por ejemplo, para calcular el Índice de Masa Corporal (IMC) a partir de peso (kg), y talla (metros). Si la variable talla está en centímetros (cm), esta variable la debes de dividir entre 100.

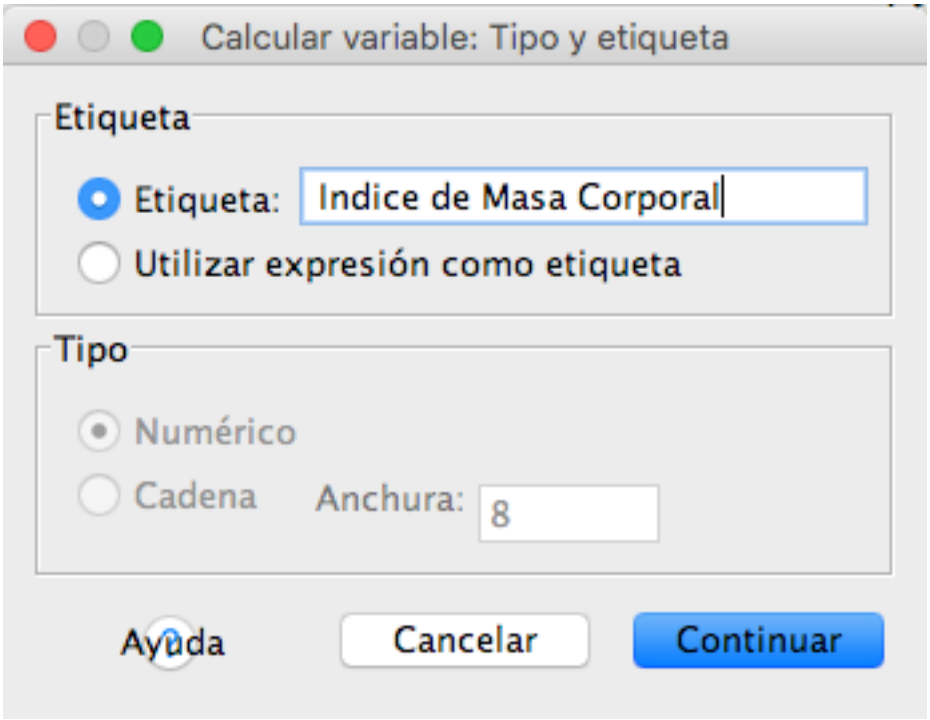


Variable destino: IMC
Expresión numérica: peso/ ((talla/100)**2)

Y pulsamos en Aceptar.

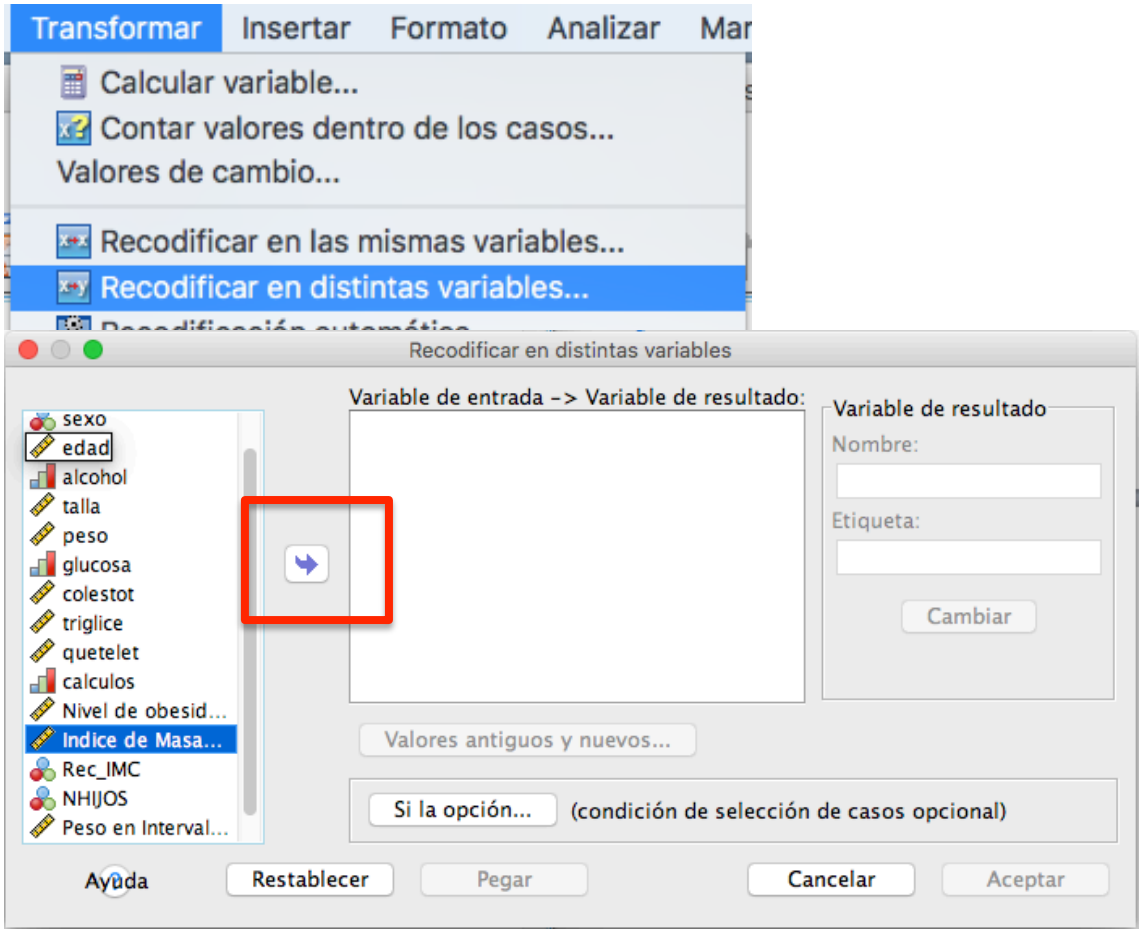


Además si pulsas en Tipo y etiqueta puedes configurar el tipo de variable, y la etiqueta.

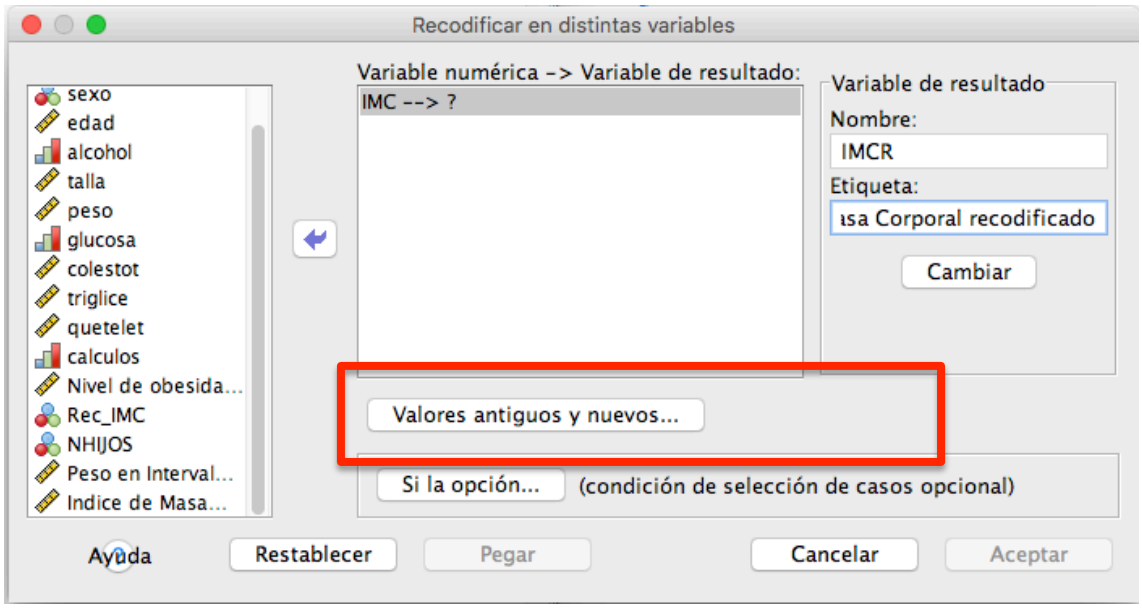


2.2 Recodificación en distintas variables.

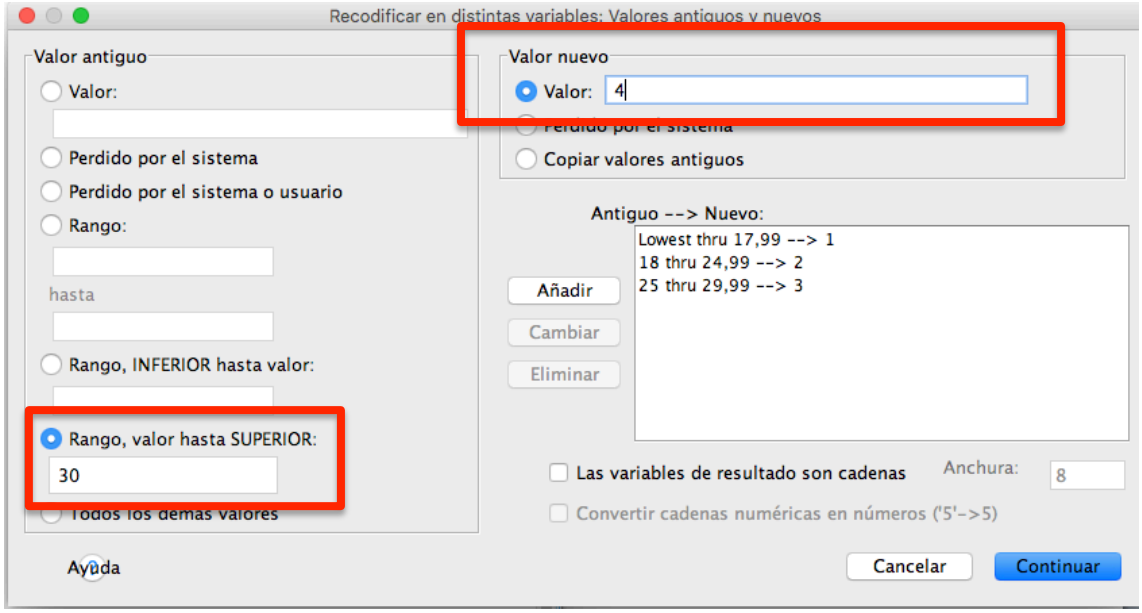
En el caso que desee recodificar los valores de la variables Índice de Masa Corporal en 4 categorías: 1:valores < 18 2: 18<=valores<25 3: 25=<valores<30 4: valores >30. Se procede de la siguiente forma:



Aquí debe de seleccionar la variable que desee recodificar y pulsar sobre la flecha azul.

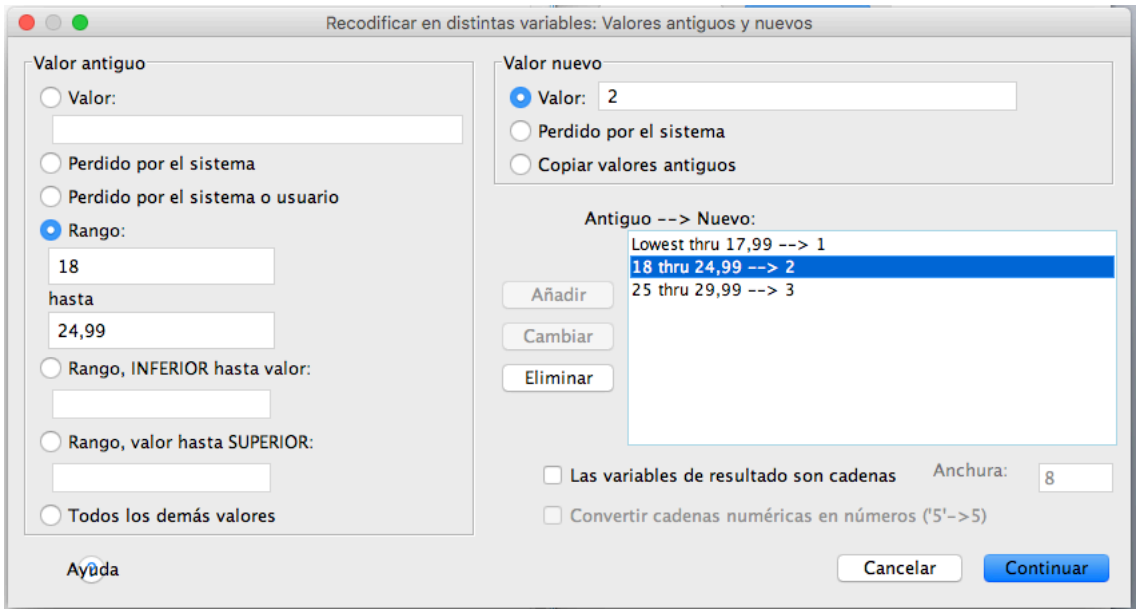


Ahora hay que pulsar en “Valores antiguos y nuevos”.



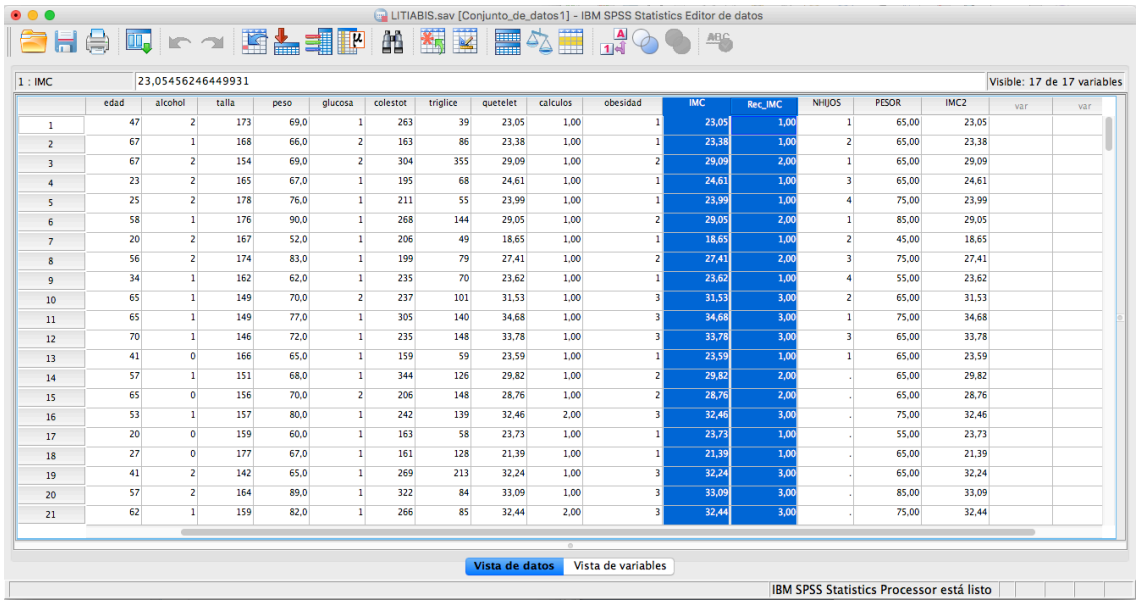
E introducir valores antiguos y valores nuevos. Por ejemplo, para valores mayores de 30, será como puedes en la captura de pantalla anterior.

Si cometes un error, puedes realizar modificaciones. Pulsas sobre la recodificación, modificas y pulsas en Cambiar.



Aquí puedes observar otra recodificación, en este caso los 28<=valores<25 toman el valor 2.

El resultado en la base de datos es el siguiente:



	edad	alcohol	talla	peso	glucosa	colestot	triglice	quetelet	calculos	obesidad	IMC	Rec_IMC	NHIJOS	PESOR	IMC2	var	var
1	47	2	173	69,0	1	263	39	23,05	1,00	1	23,05	1,00	1	65,00	23,05		
2	67	1	168	66,0	2	163	86	23,38	1,00	1	23,38	1,00	2	65,00	23,38		
3	67	2	154	69,0	2	304	355	29,09	1,00	2	29,09	2,00	1	65,00	29,09		
4	23	2	165	67,0	1	195	68	24,61	1,00	1	24,61	1,00	3	65,00	24,61		
5	25	2	178	76,0	1	211	55	23,99	1,00	1	23,99	1,00	4	75,00	23,99		
6	58	1	176	90,0	1	268	144	29,05	1,00	2	29,05	2,00	1	85,00	29,05		
7	20	2	167	52,0	1	206	49	18,65	1,00	1	18,65	1,00	2	45,00	18,65		
8	56	2	174	83,0	1	199	79	27,41	1,00	2	27,41	2,00	3	75,00	27,41		
9	34	1	162	62,0	1	235	70	23,62	1,00	1	23,62	1,00	4	55,00	23,62		
10	65	1	149	70,0	2	237	101	31,53	1,00	3	31,53	3,00	2	65,00	31,53		
11	65	1	149	77,0	1	305	140	34,68	1,00	3	34,68	3,00	1	75,00	34,68		
12	70	1	146	72,0	1	235	148	33,78	1,00	3	33,78	3,00	3	65,00	33,78		
13	41	0	166	65,0	1	159	59	23,59	1,00	1	23,59	1,00	1	65,00	23,59		
14	57	1	151	68,0	1	344	126	29,82	1,00	2	29,82	2,00	.	65,00	29,82		
15	65	0	156	70,0	2	206	148	28,76	1,00	2	28,76	2,00	.	65,00	28,76		
16	53	1	157	80,0	1	242	139	32,46	2,00	3	32,46	3,00	.	75,00	32,46		
17	20	0	159	60,0	1	163	58	23,73	1,00	1	23,73	1,00	.	55,00	23,73		
18	27	0	177	67,0	1	161	128	21,39	1,00	1	21,39	1,00	.	65,00	21,39		
19	41	2	142	65,0	1	269	213	32,24	1,00	3	32,24	3,00	.	65,00	32,24		
20	57	2	164	89,0	1	322	84	33,09	1,00	3	33,09	3,00	.	85,00	33,09		
21	62	1	159	82,0	1	266	85	32,44	2,00	3	32,44	3,00	.	75,00	32,44		

2.3 Ponderar casos

Ponderar casos proporciona a los casos diferentes ponderaciones (mediante una réplica simulada) para el análisis estadístico.

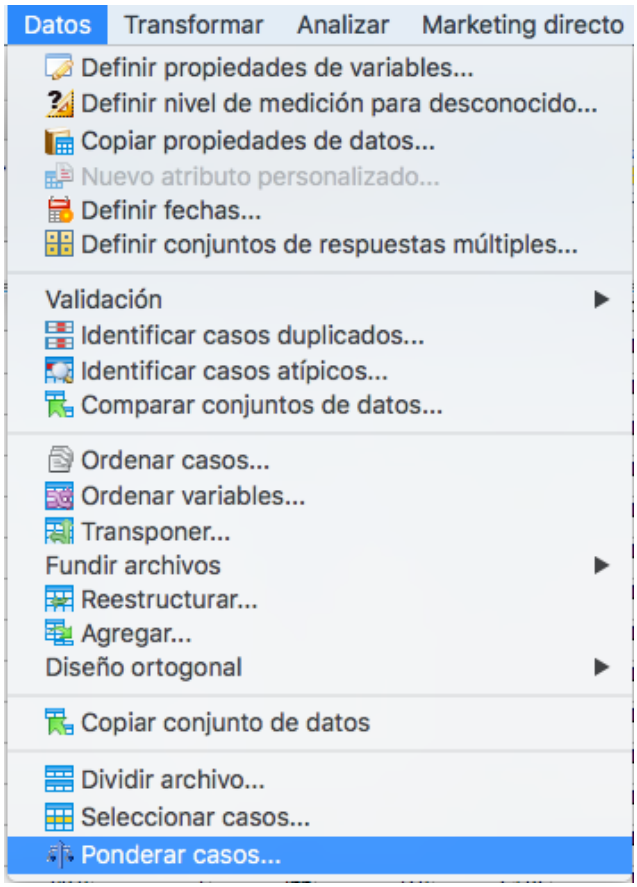
- Los valores de la variable de ponderación deben indicar el número de observaciones representadas por casos únicos en el archivo de datos.

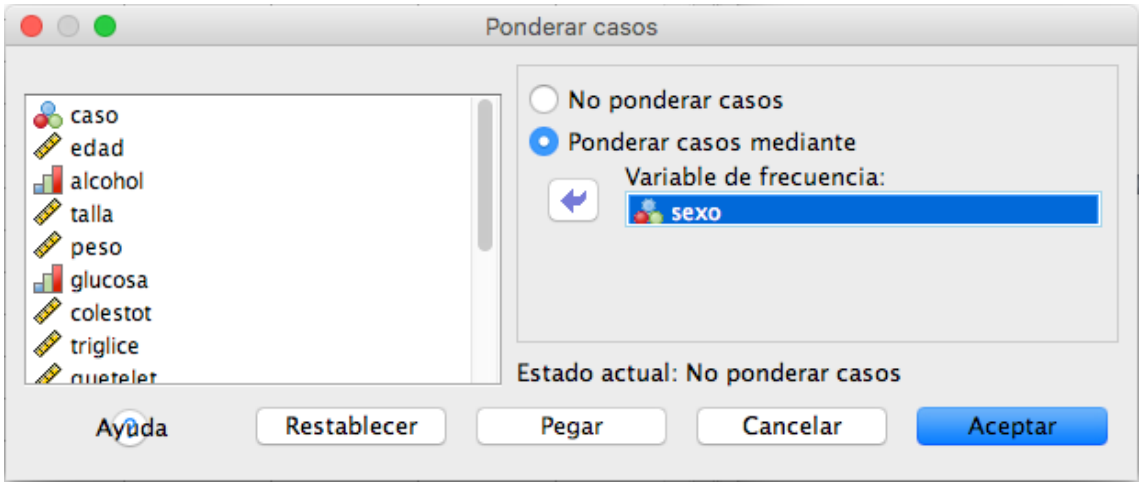
- Los casos con valores perdidos, negativos o cero para la variable de ponderación se excluyen del análisis.

- Los valores fraccionarios son válidos y algunos procedimientos, como Frecuencias, Tablas de contingencia y Tablas personalizadas, utilizan valores de ponderación fraccionaria. Sin embargo, la mayoría de los procedimientos consideran la ponderación de variables una ponderación de replicación y simplemente redondean las ponderaciones fraccionarias al número entero más cercano. Algunos procedimientos ignoran por completo la variable de ponderación, y esta limitación se indica en la documentación específica del procedimiento.

Ponderaciones en las tablas de contingencia. El procedimiento Tablas de contingencia cuenta con diversas opciones para el tratamiento de ponderaciones de los casos.

Ponderaciones en los diagramas de dispersión y los histogramas. Los diagramas de dispersión y los histogramas tienen una opción para activar y desactivar las ponderaciones de los casos, pero dicha opción no afecta a los casos que tienen un valor negativo, un valor 0 o un valor perdido para la variable de ponderación. Estos casos permanecen excluidos del gráfico incluso si se desactiva la ponderación desde el gráfico.





Seleccionas la variable a ponderar y la introduces en “Ponderar caso mediante”.

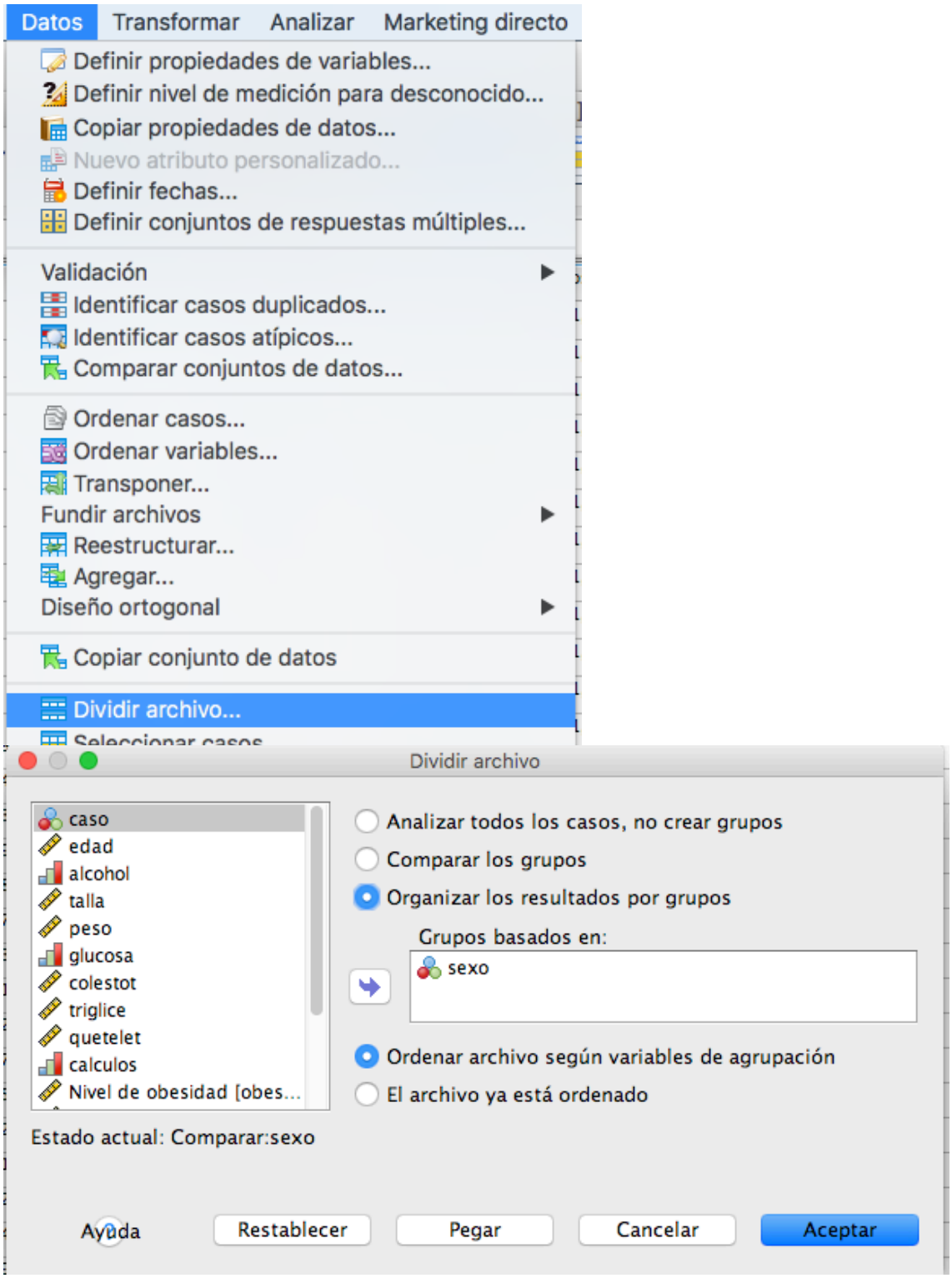
2.4 Dividir/Segmentar archivo

Segmentar archivo divide el archivo de datos en distintos grupos para el análisis basándose en los valores de una o más variables de agrupación. Si selecciona varias variables de agrupación, los casos se agruparán por variable dentro de las categorías de la variable anterior de la lista Grupos basados en. Por ejemplo, si selecciona *sexo* como la primera variable de agrupación y *minoría* como la segunda, los casos se agruparán por minorías dentro de cada categoría de sexo.

- Es posible especificar hasta ocho variables de agrupación.
- Cada ocho bytes de una variable de cadena larga (variables de cadena que superan los ocho bytes) cuenta como una variable hasta llegar al límite de ocho variables de agrupación.
- Los casos deben ordenarse según los valores de las variables de agrupación, en el mismo orden en el que aparecen las variables en la lista Grupos basados en. Si el archivo de datos todavía no está ordenado, seleccione **Ordenar archivo según variables de agrupación**.

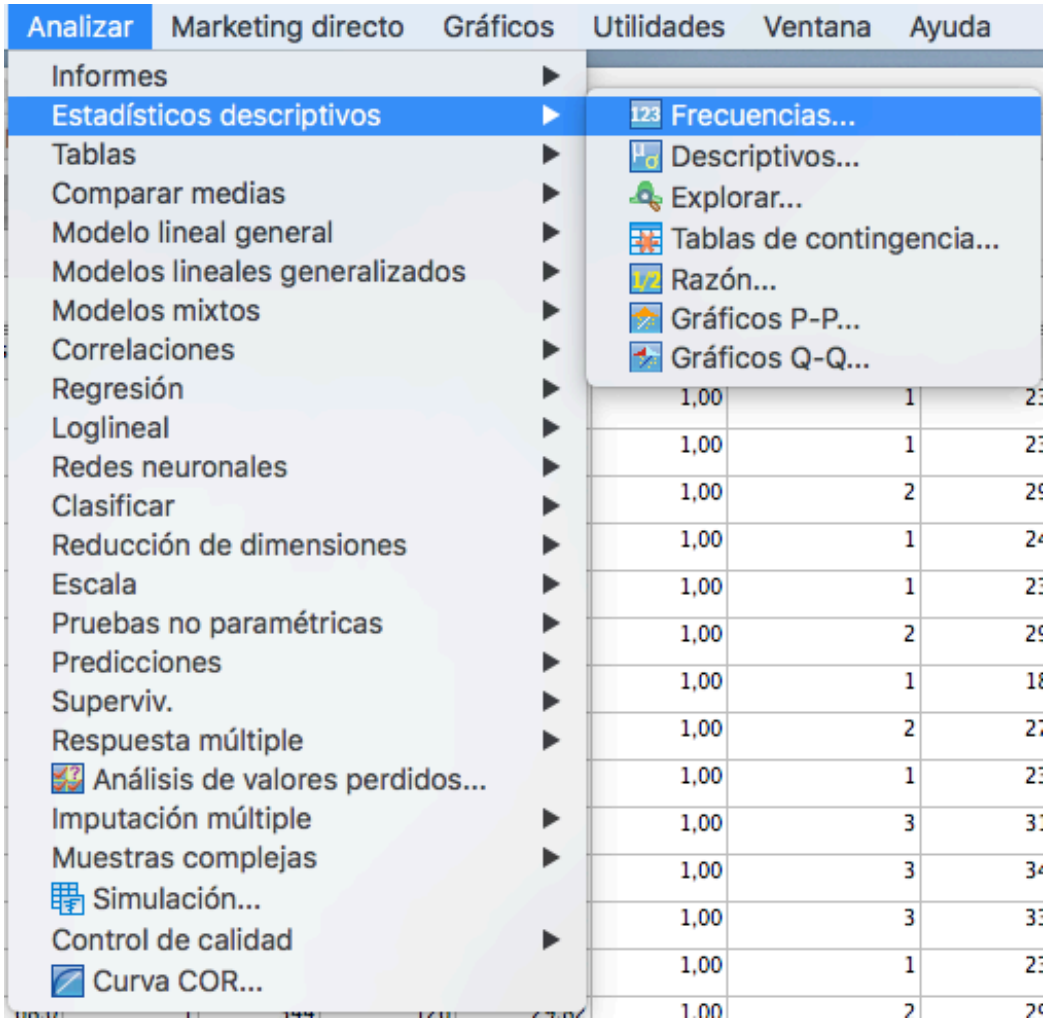
Comparar los grupos. Los grupos de segmentación del archivo se presentan juntos para poder compararlos. Para las tablas pivote se crea una sola tabla y cada variable de segmentación del archivo puede desplazarse entre las dimensiones de la tabla. En el caso de los gráficos se crea un gráfico diferente para cada grupo y se muestran juntos en el Visor.

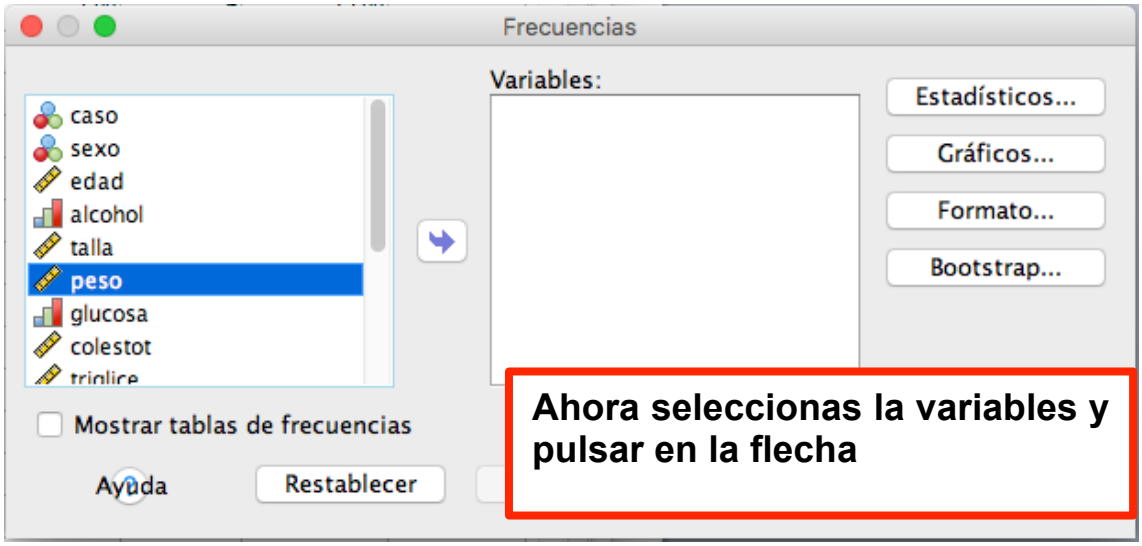
Organizar los resultados por grupos. Los resultados de cada procedimiento se muestran por separado para cada grupo de segmentación del archivo.

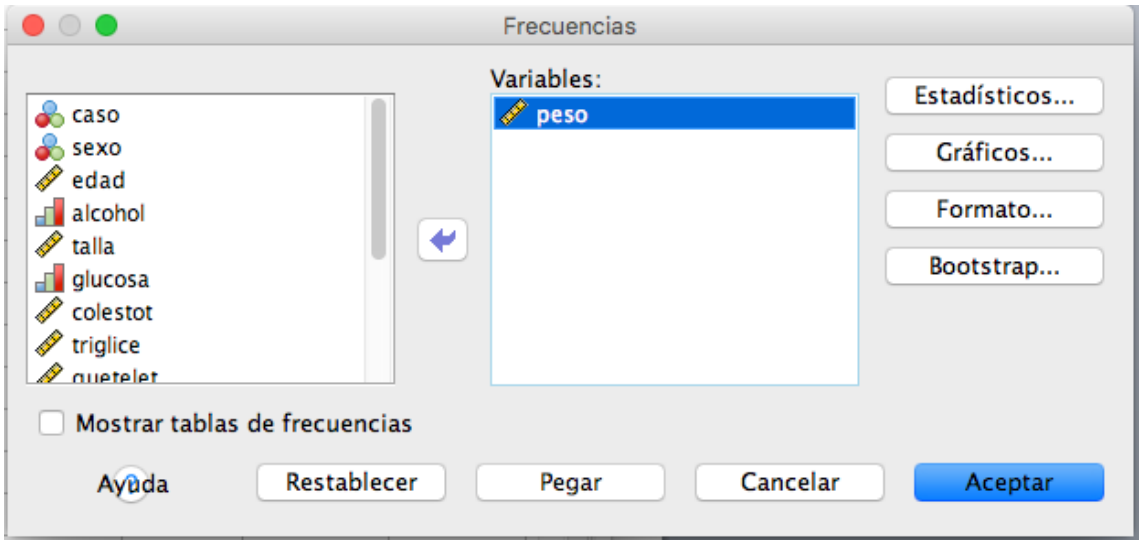

Datos
Transformar
Analizar
Marketing directo
Definir propiedades de variables...
Definir nivel de medición para desconocido...
Copiar propiedades de datos...
Nuevo atributo personalizado...
Definir fechas...
Definir conjuntos de respuestas múltiples...
Validación
Identificar casos duplicados...
Identificar casos atípicos...
Comparar conjuntos de datos...
Ordenar casos...
Ordenar variables...
Transponer...
Fundir archivos
Reestructurar...
Agregar...
Diseño ortogonal
Copiar conjunto de datos
Dividir archivo...
Dividir archivo
caso
edad
alcohol
talla
peso
glucosa
colestot
triglice
quetelet
calculos
Nivel de obesidad [obes...
Analizar todos los casos, no crear grupos
Comparar los grupos
Organizar los resultados por grupos
Grupos basados en:
sexo
Ordenar archivo según variables de agrupación
El archivo ya está ordenado
Estado actual: Comparar:sexo
Ayuda
Restablecer
Pegar
Cancelar
Aceptar

3. Análisis estadístico

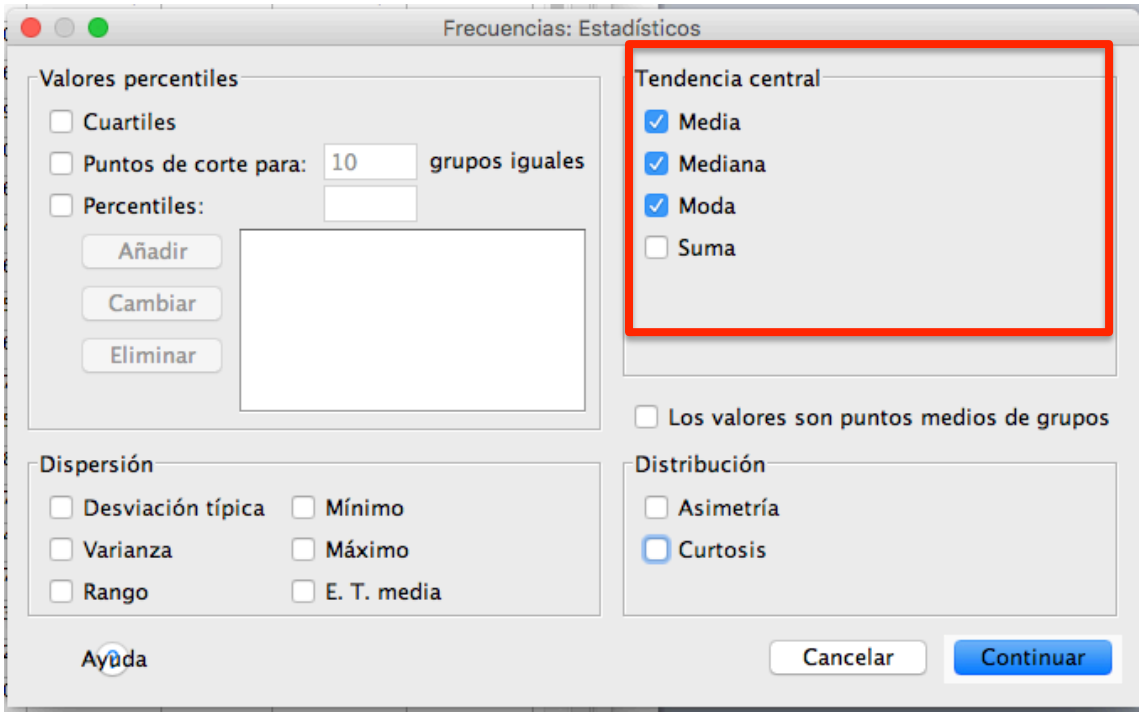
3.1 Medidas de tendencia central



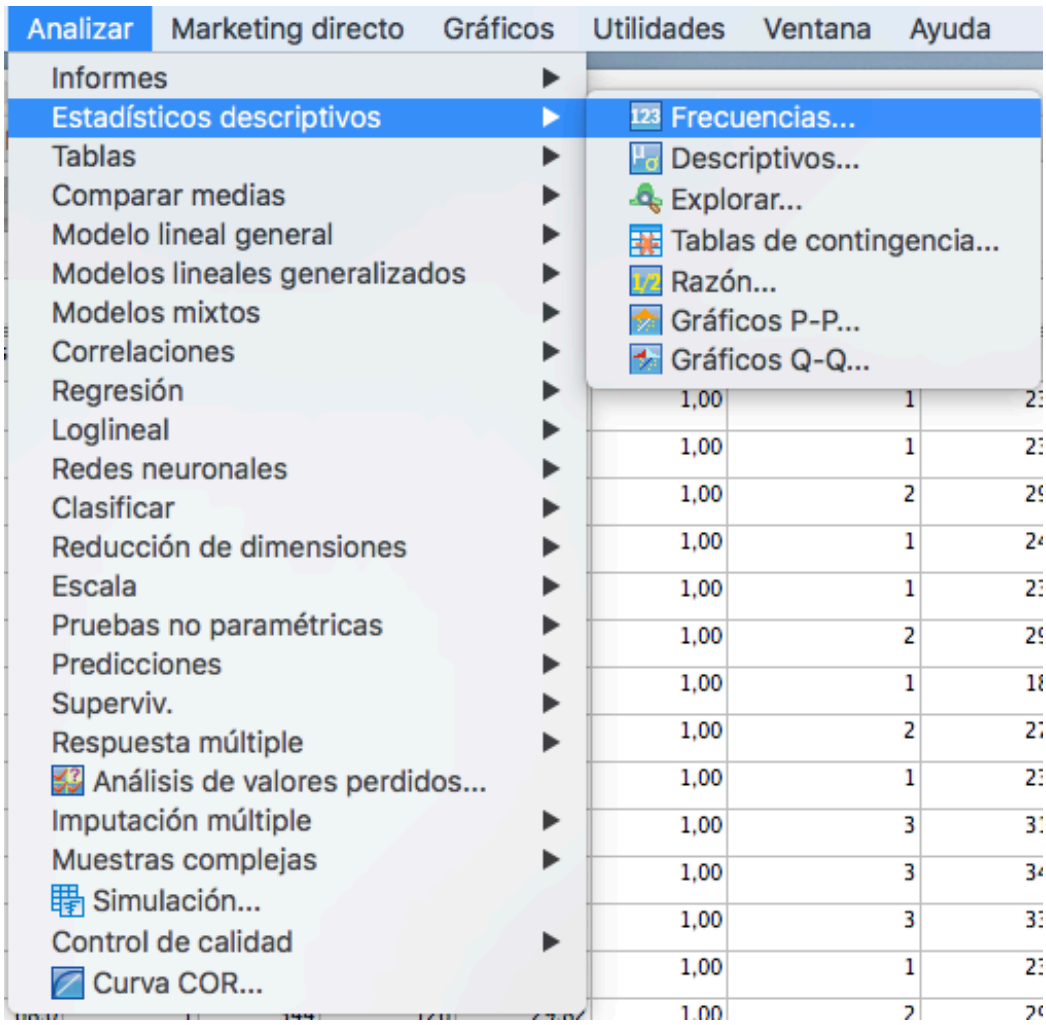


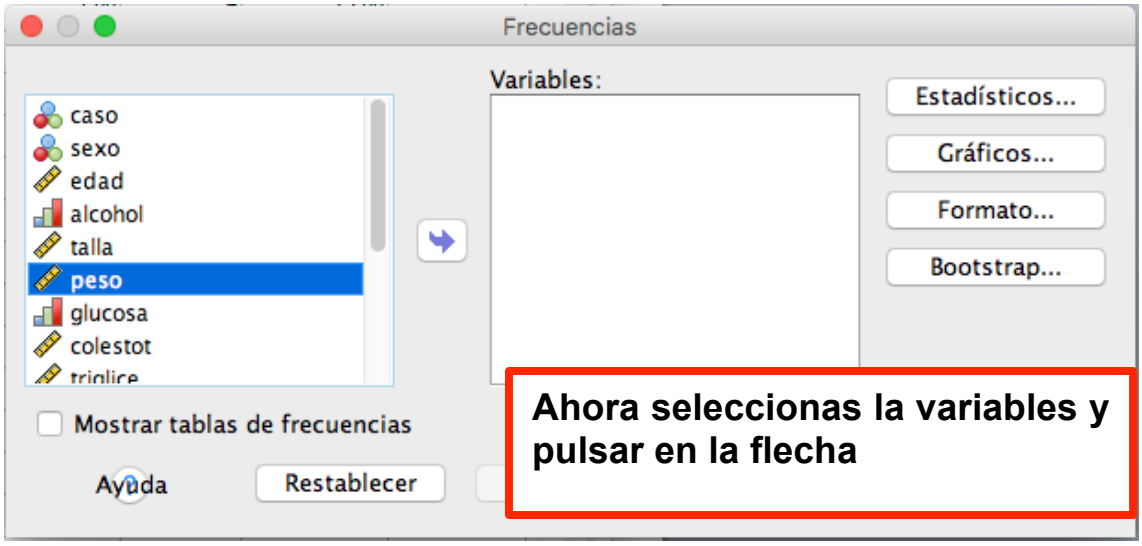


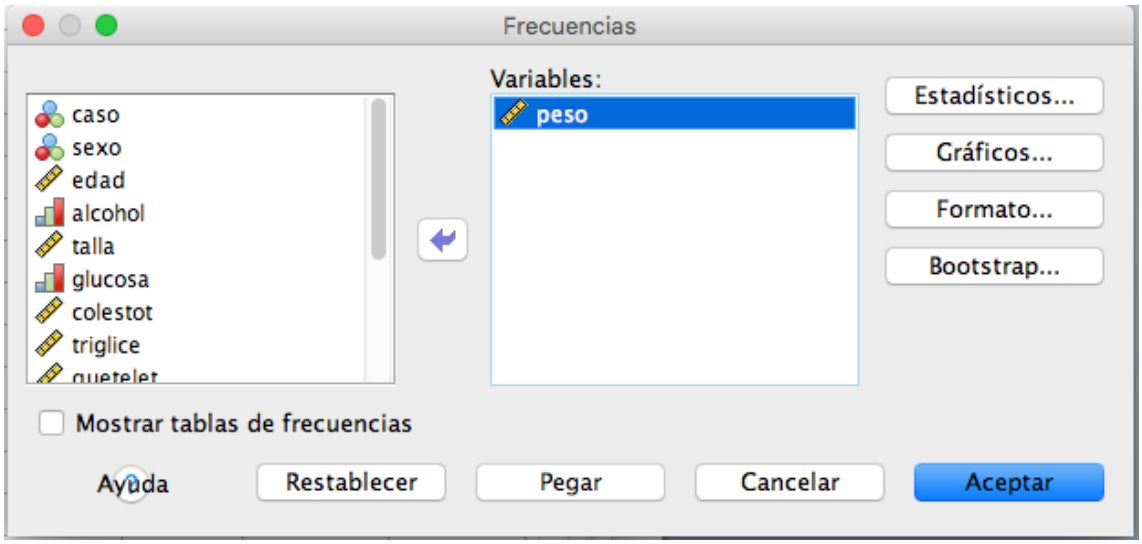
Ahora debes de pulsar en Estadísticos y seleccionara las medidas de tendencia central.



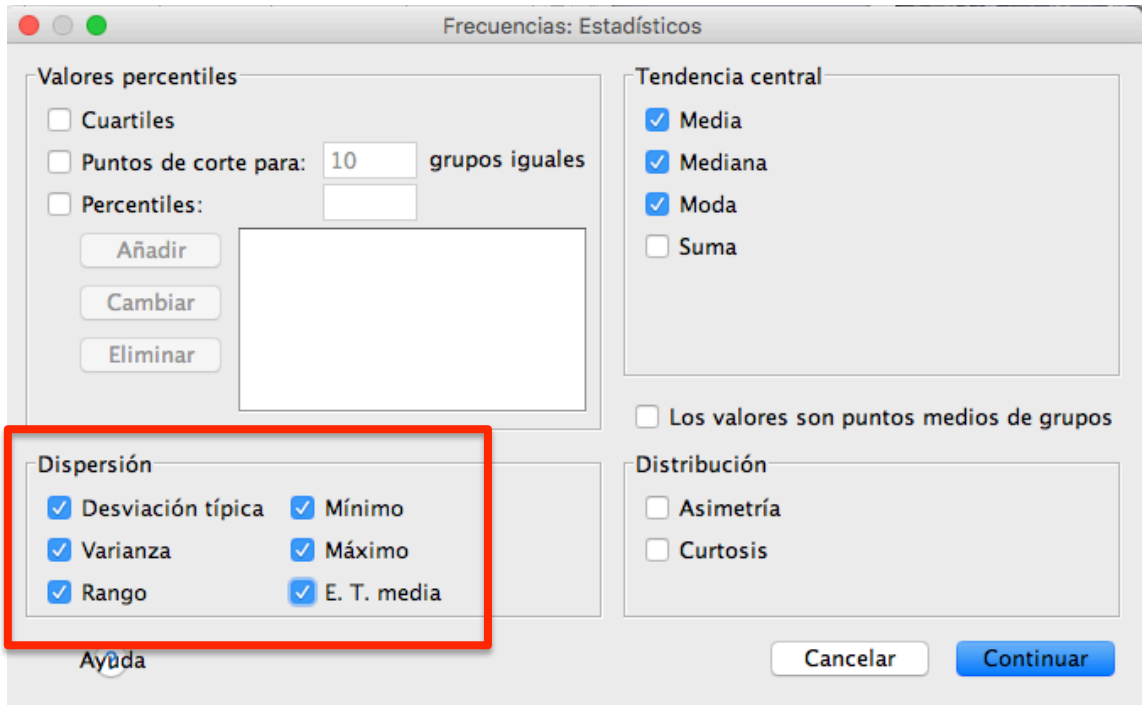
3.2 Medidas de dispersión



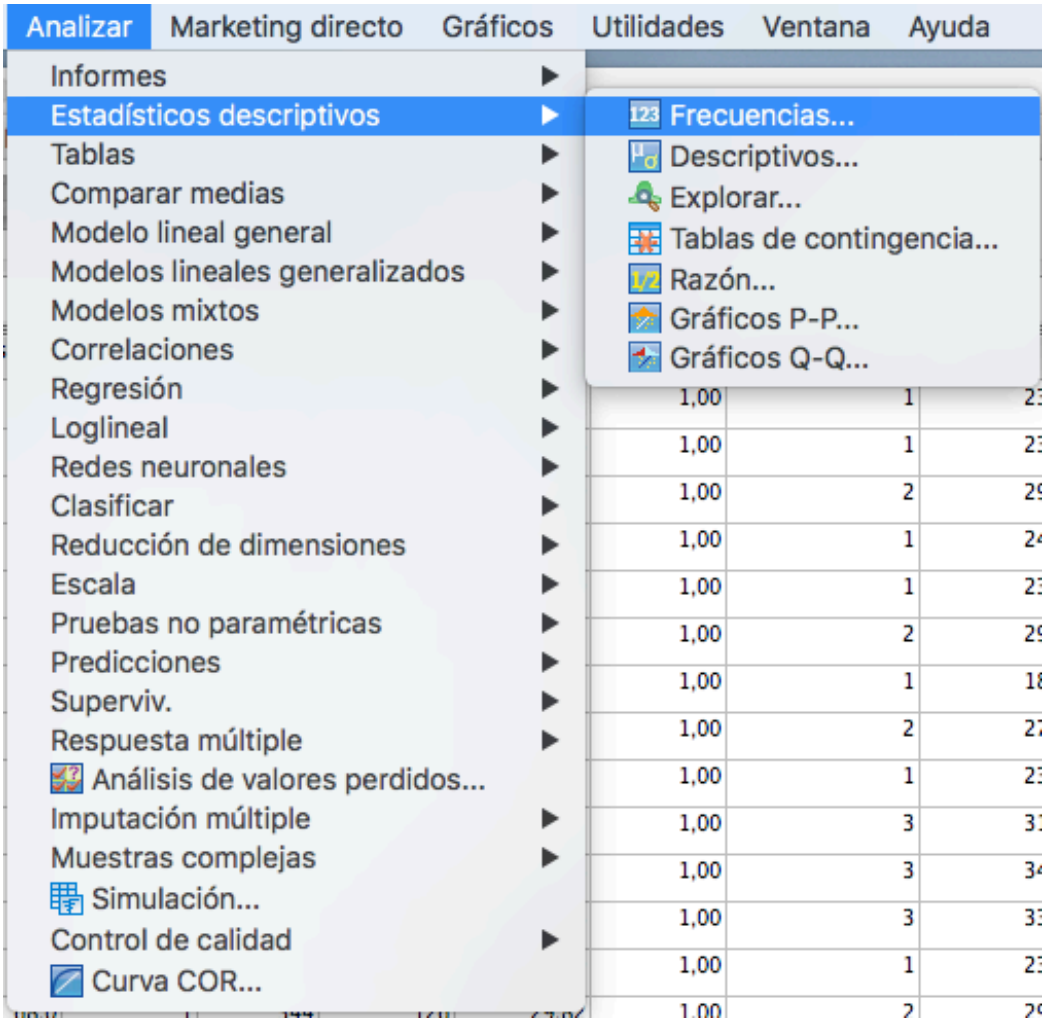


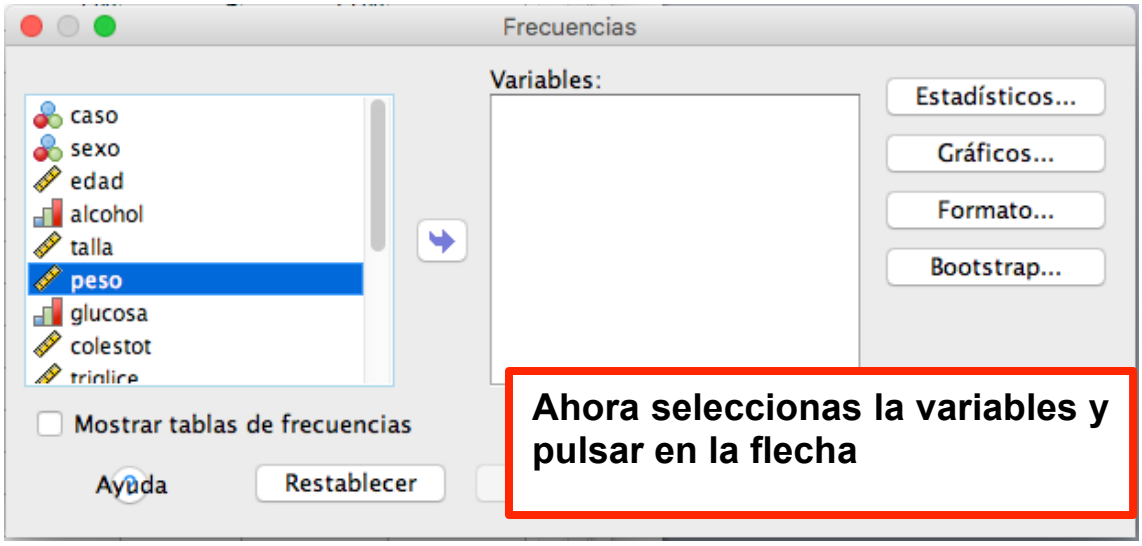


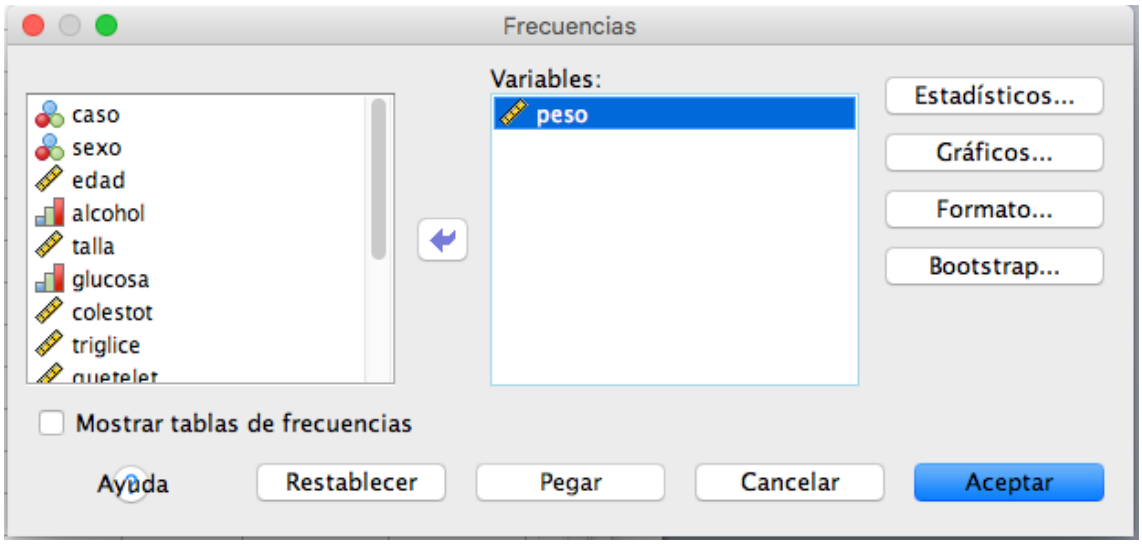
Ahora debes de pulsar en Estadísticos y seleccionara las medidas de dispersión.



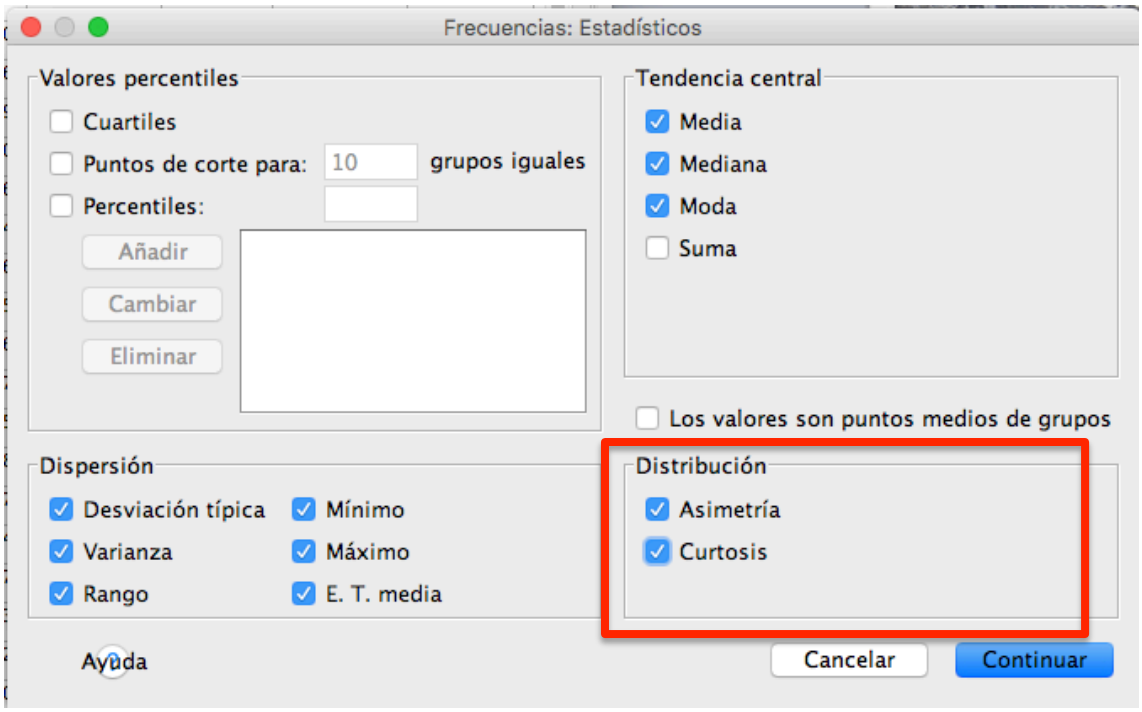
3.3 Medidas de forma



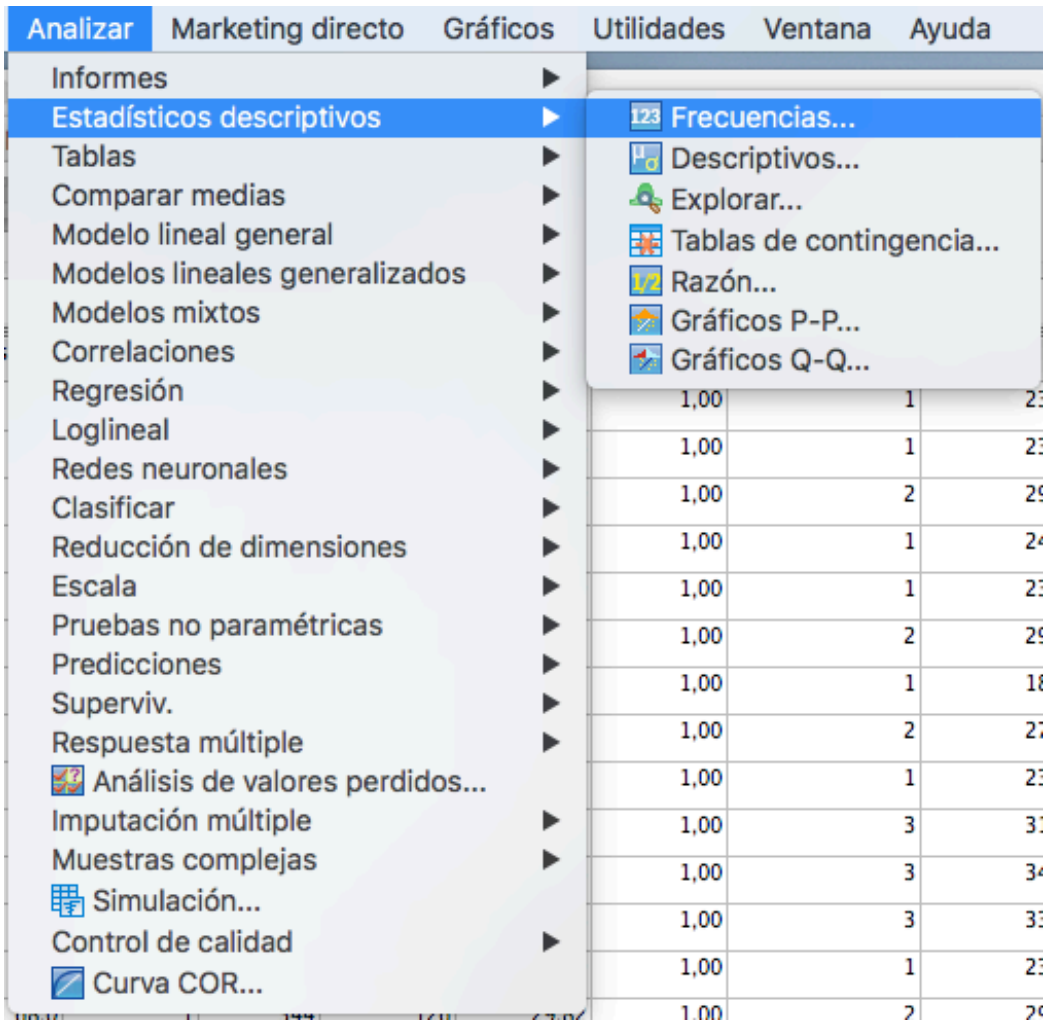


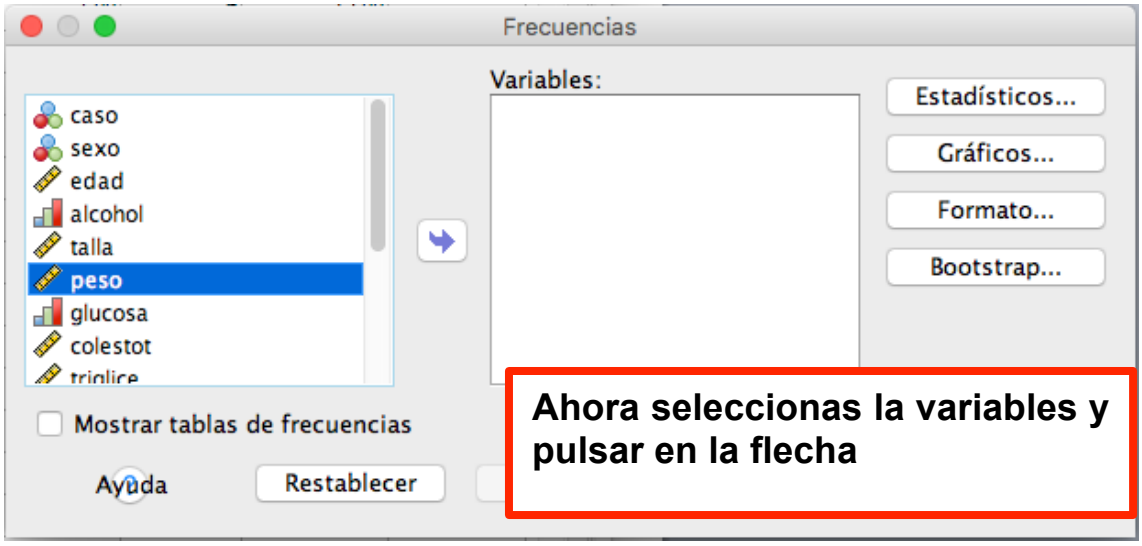


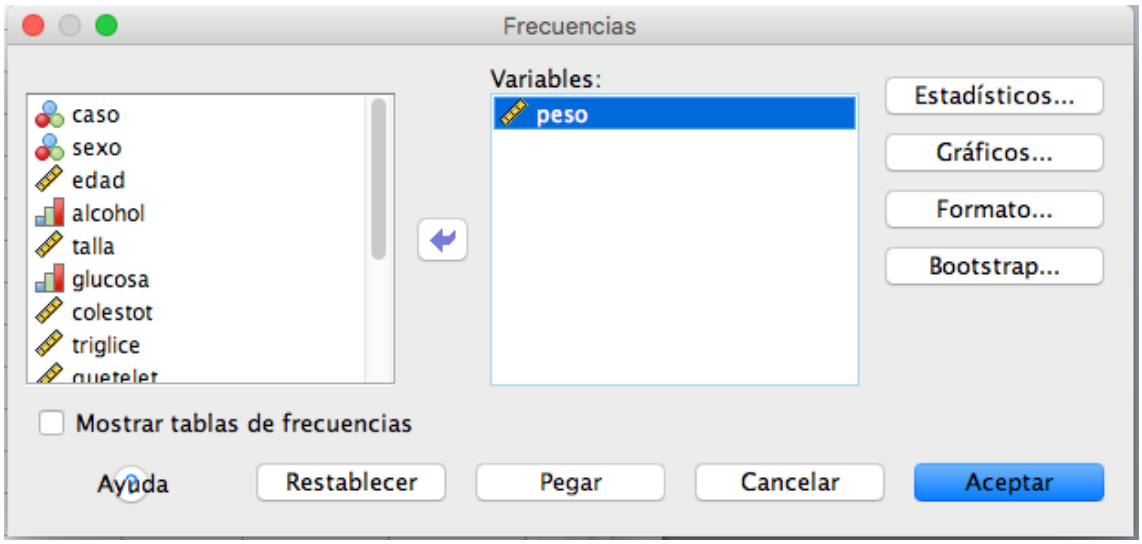
Ahora debes de pulsar en Estadísticos y seleccionar las medidas de forma.



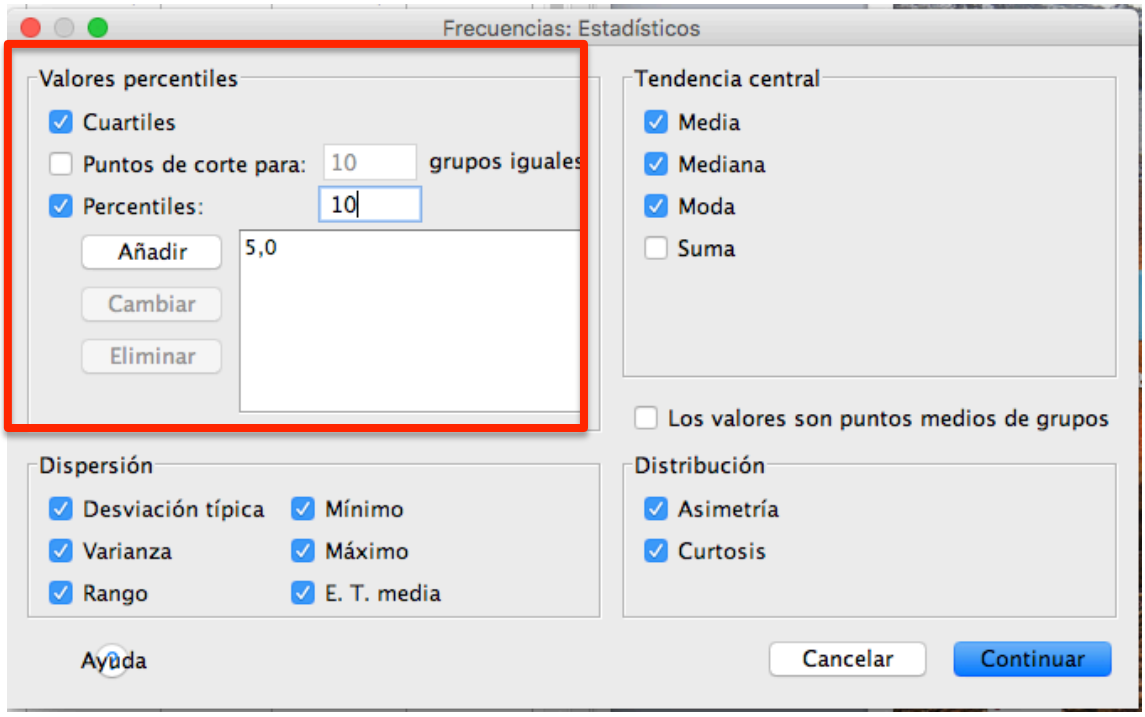
3.4 Medidas de posición







Ahora debes de pulsar en Estadísticos y seleccionara las medidas de forma.



Si quieres calcular percentiles, debes de seleccionar Percentiles, rellenar el rectángulo y pulsar en añadir.

3.5 Resultados

Los estadísticos de frecuencia, una vez en el visor de resultados se pueden copiar y pegar en cualquier editor de textos, y modificar. Debes de seleccionar la tabla en el fichero de resultados y copiar en el editor.

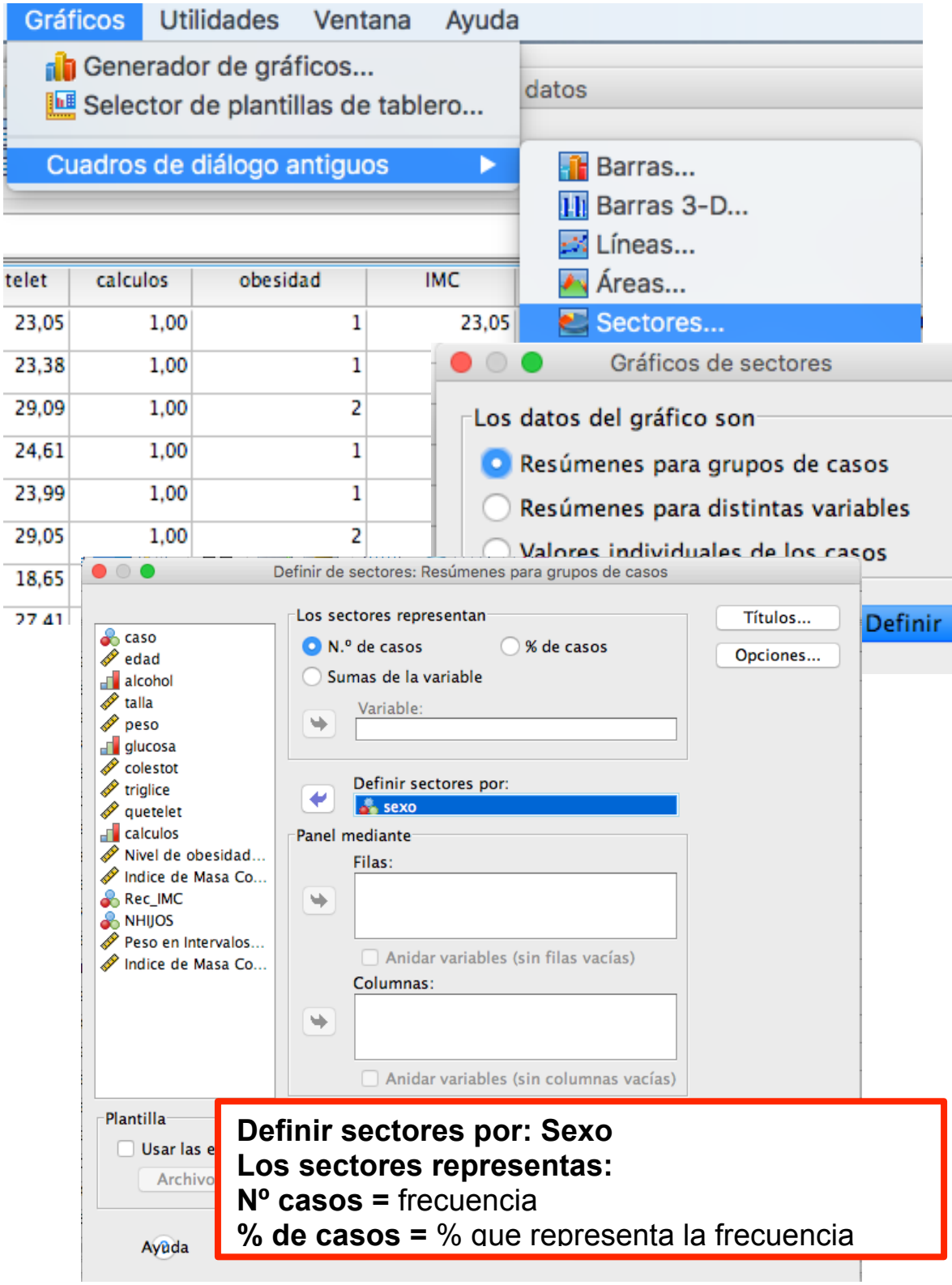
peso

N	Valid	200
	Missing	0
Mean		70,210
Std. Error of Mean		,7999
Median		69,000
Mode		65,0
Std. Deviation		11,3125
Variance		127,973
Skewness		,247
Std. Error of Skewness		,172
Kurtosis		-,314
Std. Error of Kurtosis		,342
Range		57,0
Minimum		44,0
Maximum		101,0
Percentiles	5	52,000
	10	55,000
	25	63,000
	50	69,000
	75	78,000

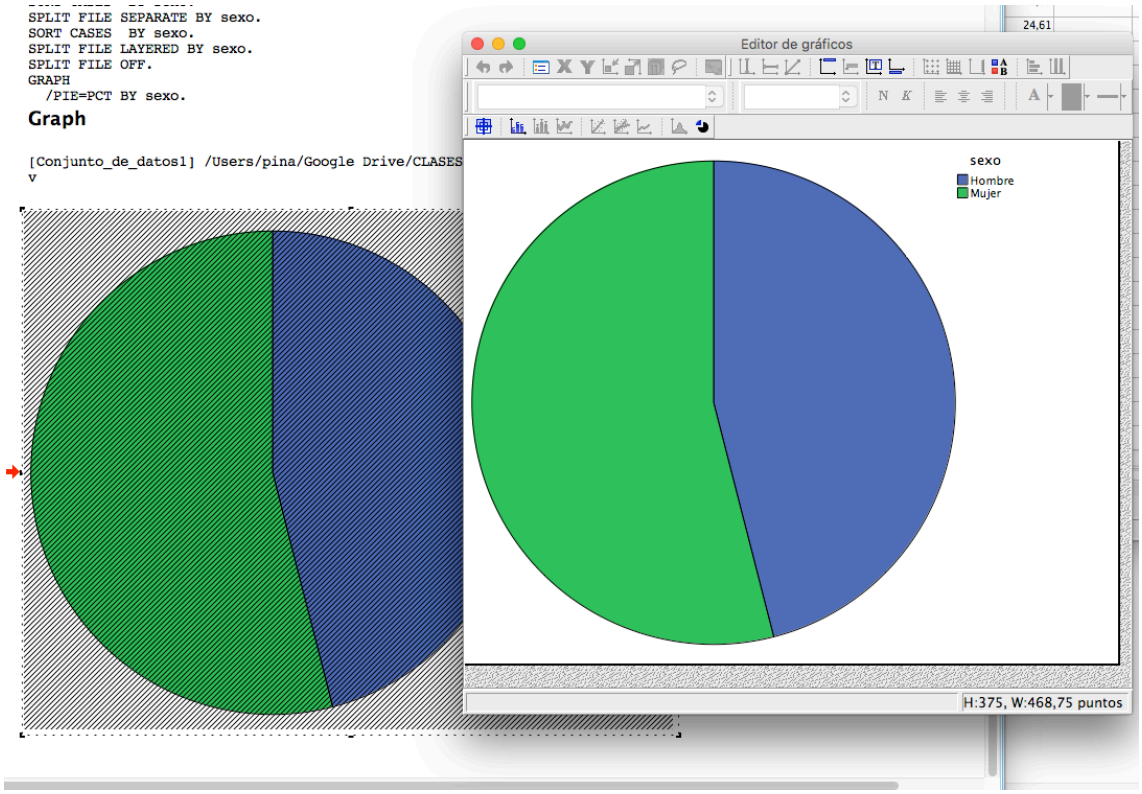
4. Gráficos

4.1 Gráficos de sectores.

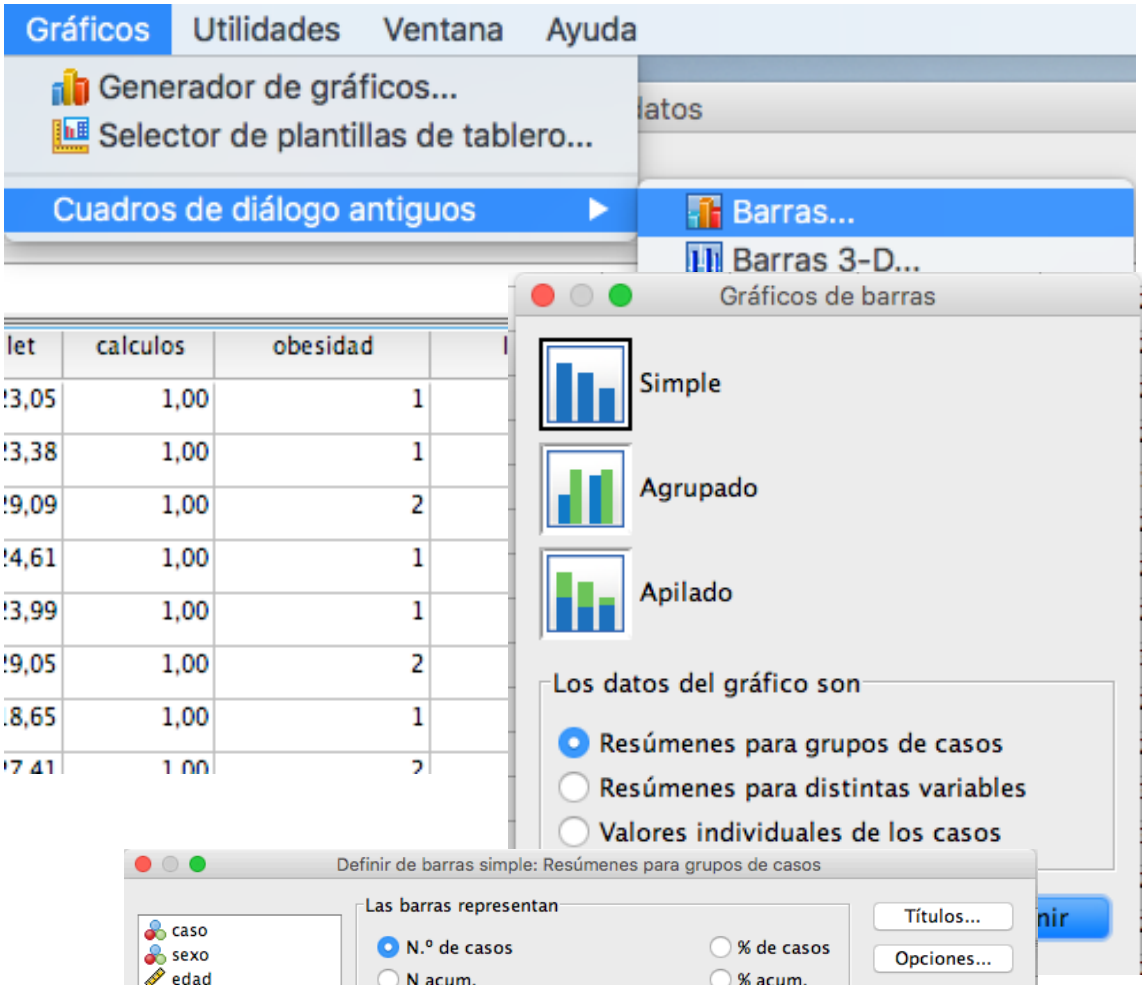
Para variables de tipo cualitativo (sexo, color de ojos, grupo sanguíneo, etc…) o de tipo categórico.

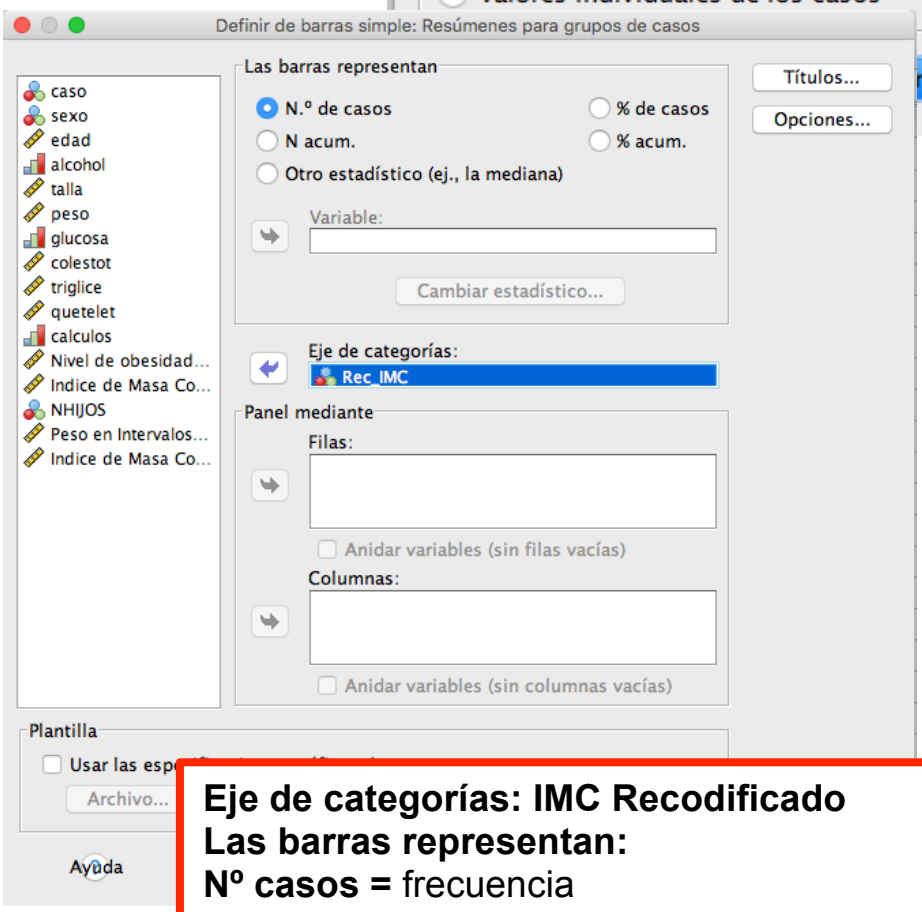


Pulsando dos clics sobre el gráfico se puede editar, y por tanto puedes modificar los colores, insertar etiquetas, etc…



4.2 Gráficos de barras.

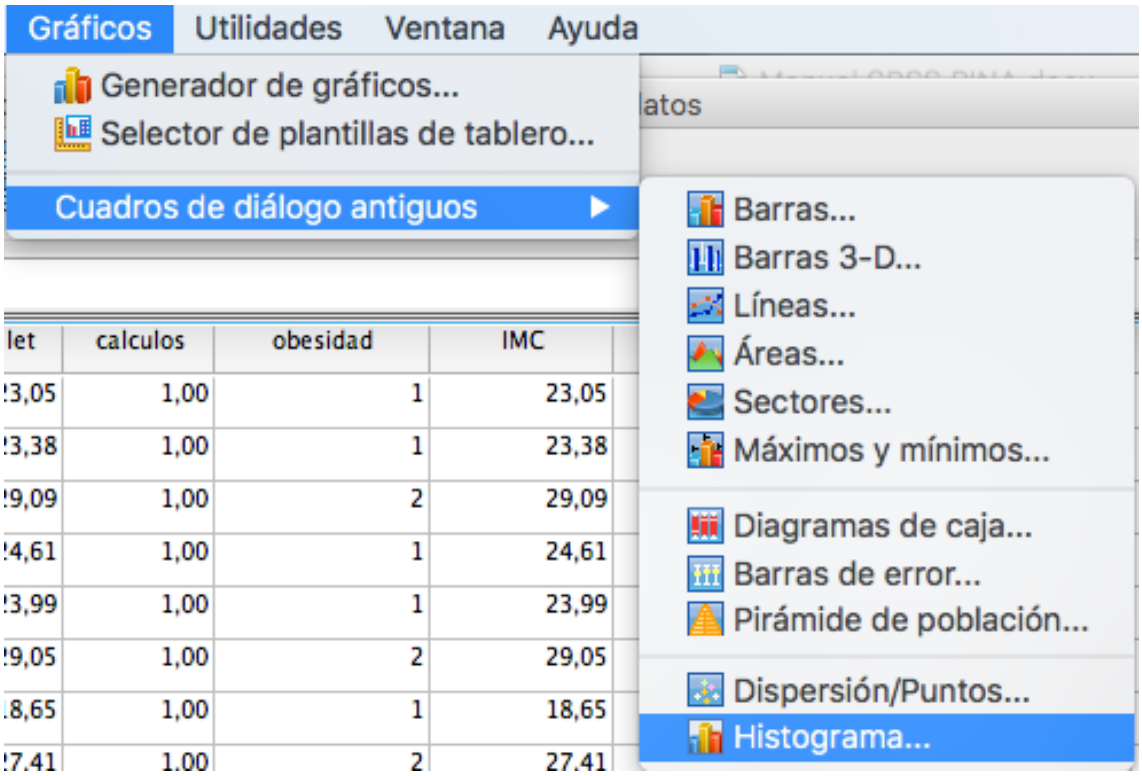


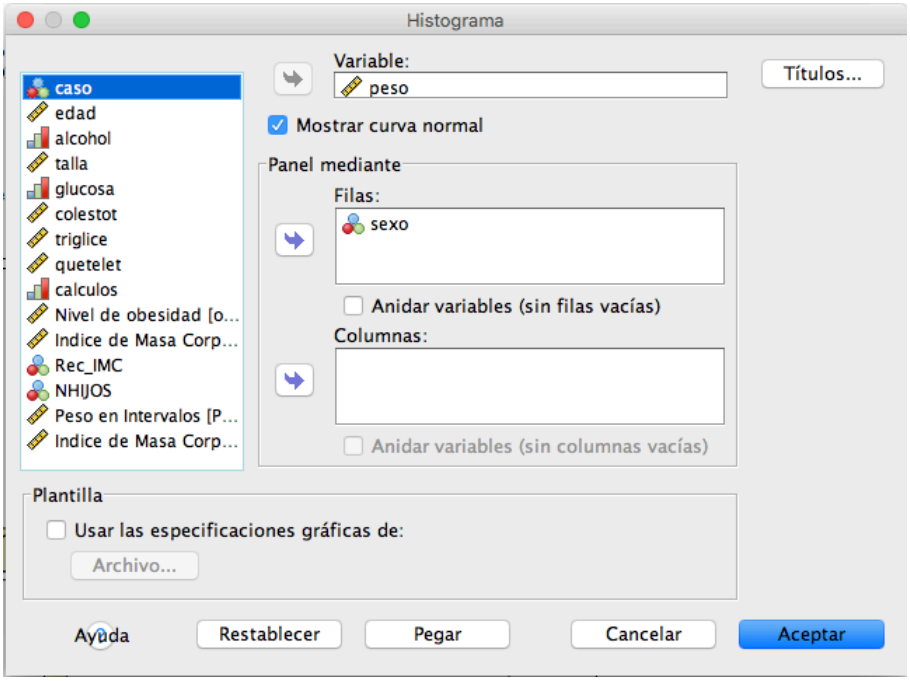


Eje de categorías: IMC Recodificado
Las barras representan:
Nº casos = frecuencia
% de casos = % que representa la frecuencia

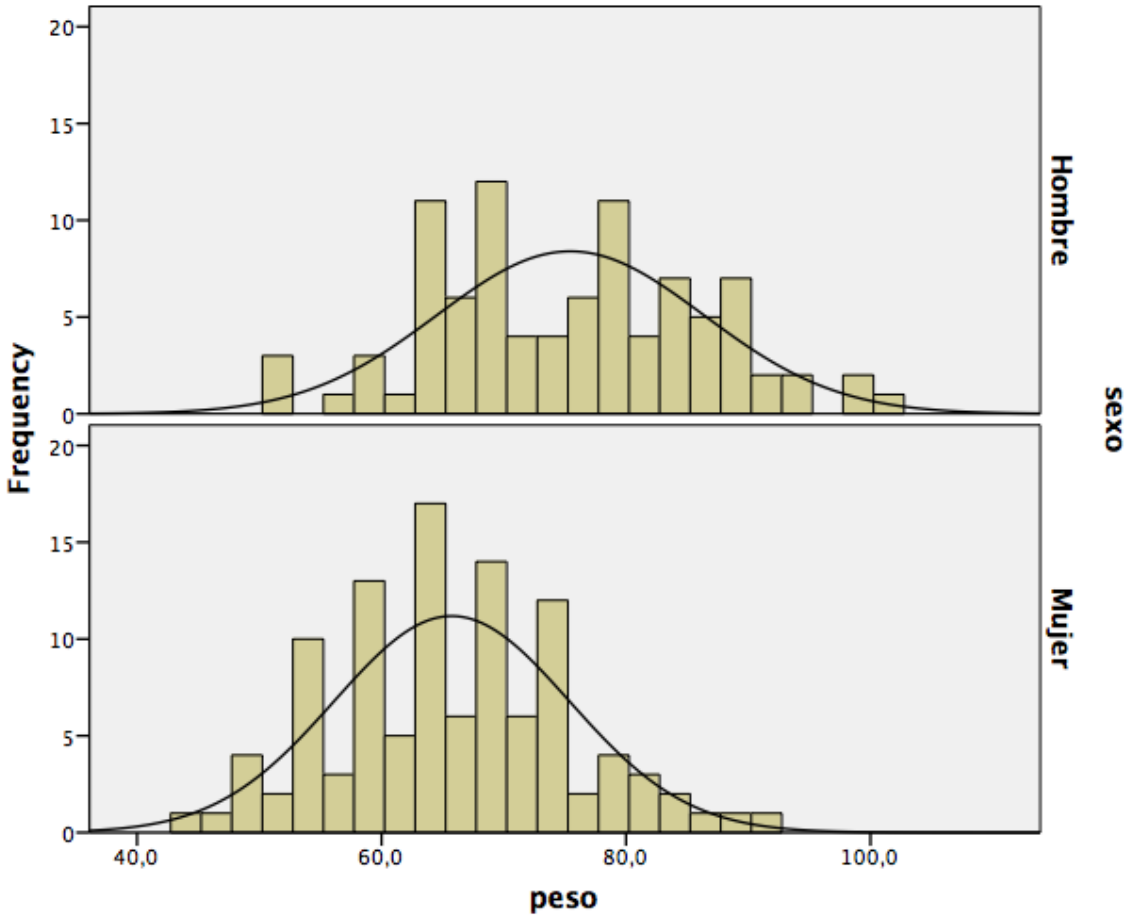
4.4 Histograma.

Es una representación gráfica de una variable, cuantitativa y continua (peso, talla) en forma de barras, donde la superficie de cada barra es proporcional a la frecuencia de los valores representados. Sirven para obtener una "primera vista" general, o panorama, de la distribución de la población, o de la muestra, respecto a una característica.



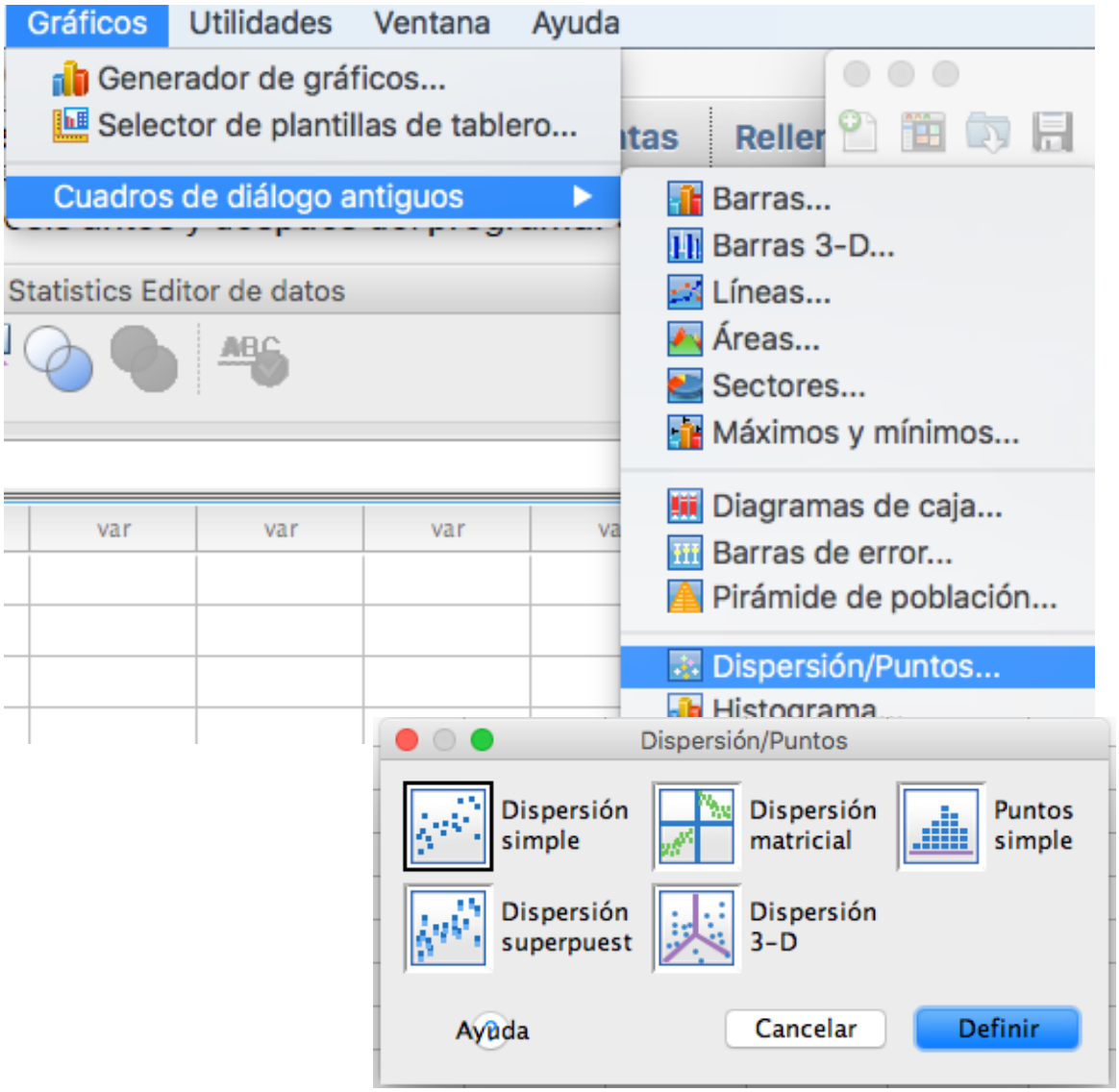


Y queda:



4.5 Gráficos de dispersión.

Un diagrama de dispersión o gráfica de dispersión o gráfico de dispersión es un tipo de diagrama matemático que utiliza las coordenadas cartesianas para mostrar los valores de dos variables para un conjunto de datos.



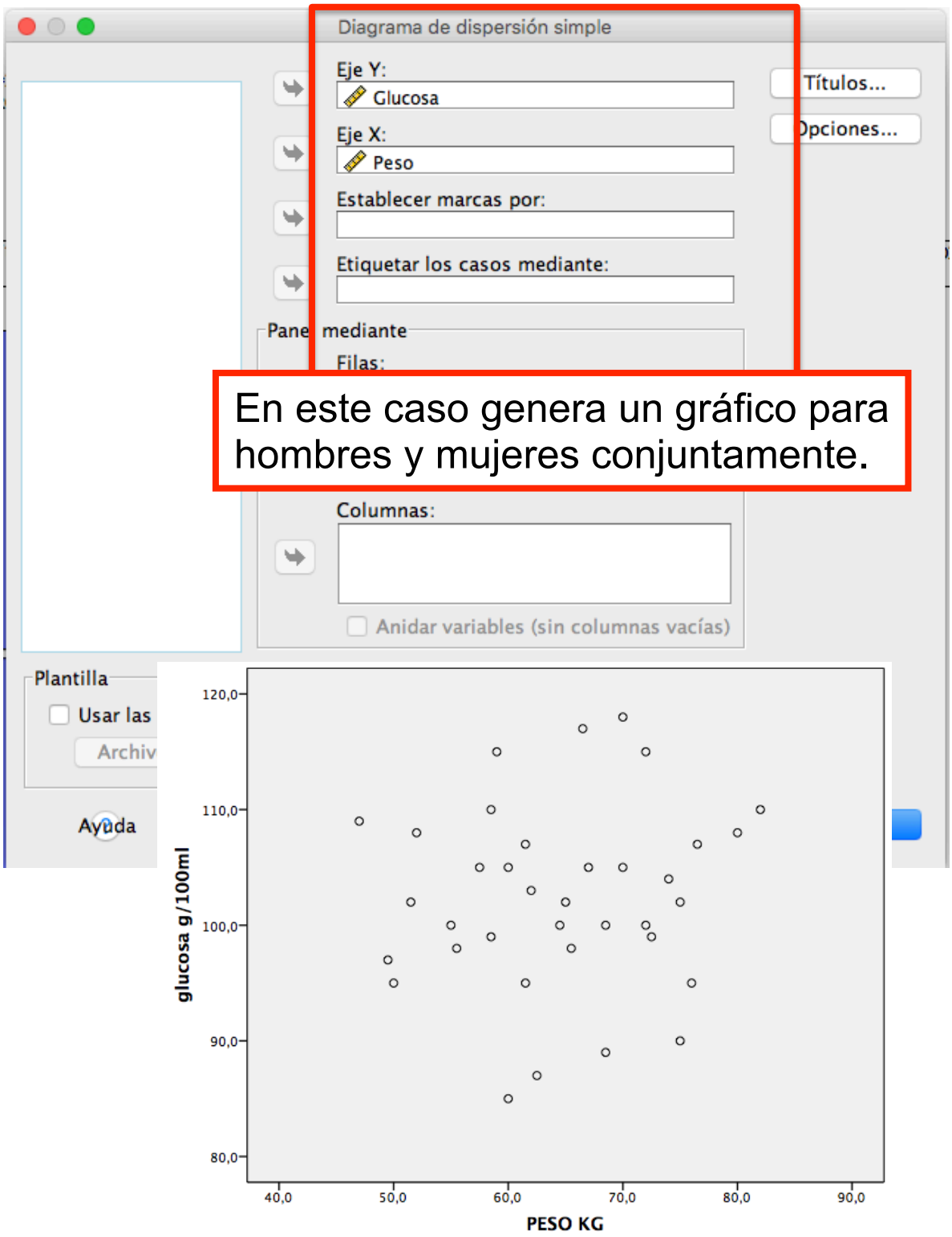
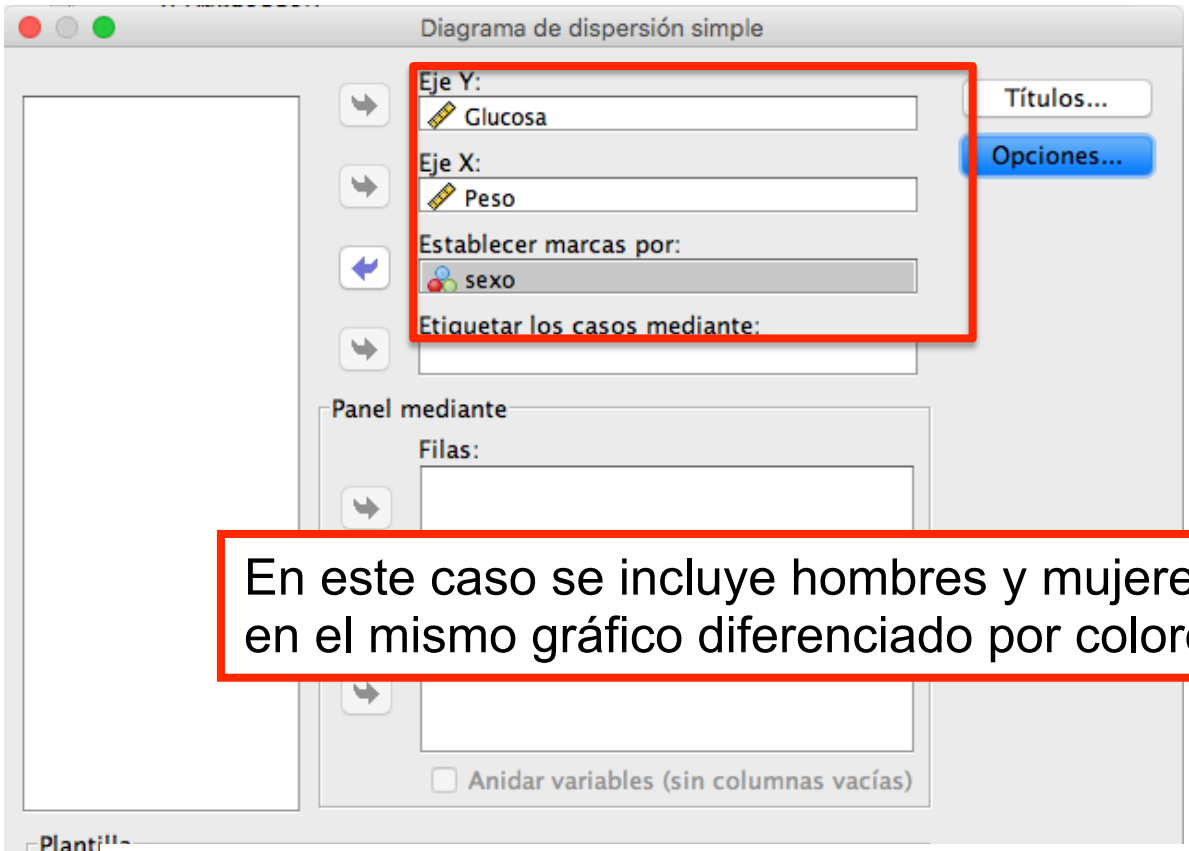
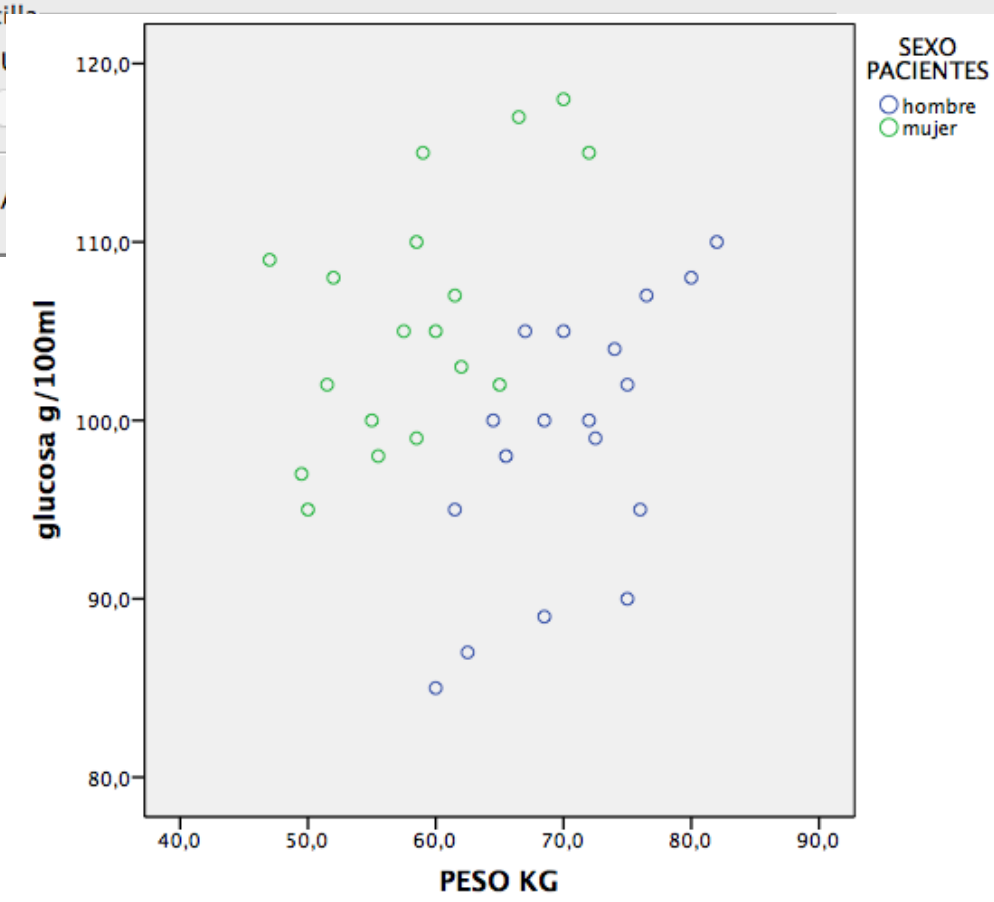


Diagrama de dispersión simple
Eje Y:
Glucosa
Eje X:
Peso
Establecer marcas por:
Etiquetar los casos mediante:
Títulos...
Opciones...
Pane mediante
Filas:
Columnas:
Anidar variables (sin columnas vacías)
Plantilla
Usar las
Archiv
Ayuda
En este caso genera un gráfico para hombres y mujeres conjuntamente.
glucosa g/100ml
120,0
110,0
100,0
90,0
80,0
40,0
50,0
60,0
70,0
80,0
90,0
PESO KG



En este caso se incluye hombres y mujeres en el mismo gráfico diferenciado por colores.



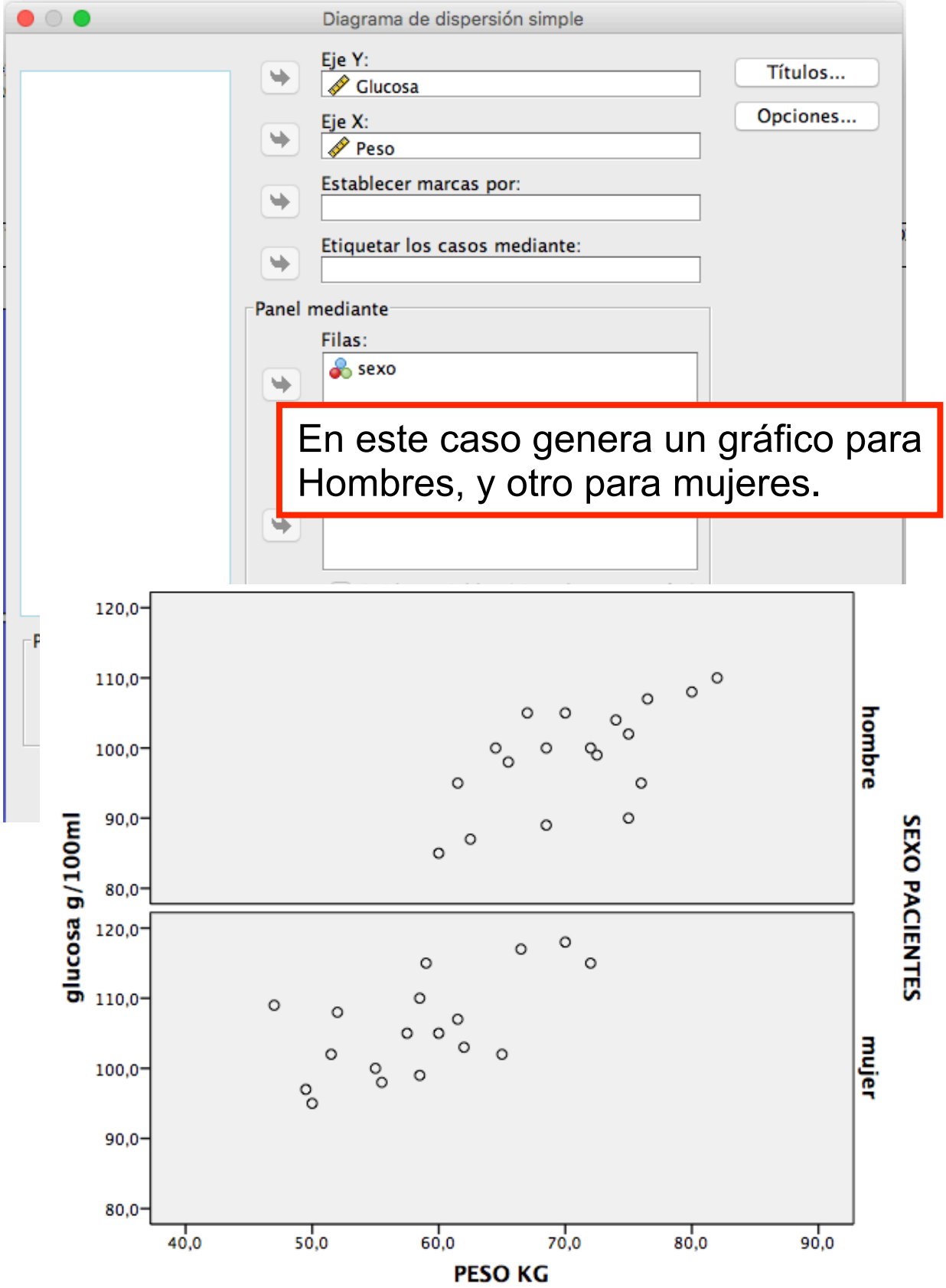
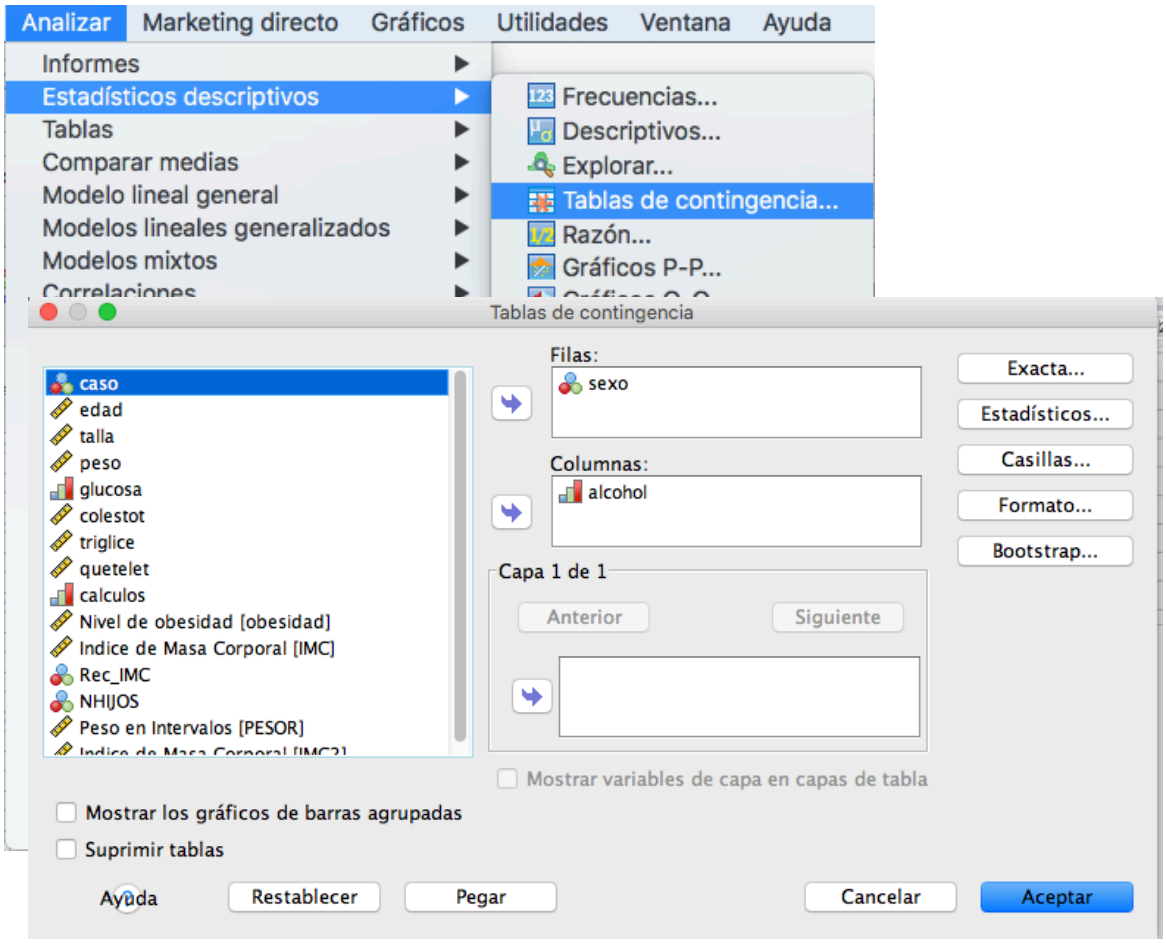


Diagrama de dispersión simple
Eje Y:
Glucosa
Eje X:
Peso
Establecer marcas por:
Etiquetar los casos mediante:
Títulos...
Opciones...
Panel mediante
Filas:
sexo
En este caso genera un gráfico para Hombres, y otro para mujeres.
glucosa g/100ml
120,0
110,0
100,0
90,0
80,0
hombre
mujer
SEXO PACIENTES
40,0
50,0
60,0
70,0
80,0
90,0
PESO KG

5. Análisis de variables categóricas (tablas de contingencia.

Las variables deben de ser cualitativas, o variables cuantitativas discretas, o variables cuantitativas recodificas por intervalos.



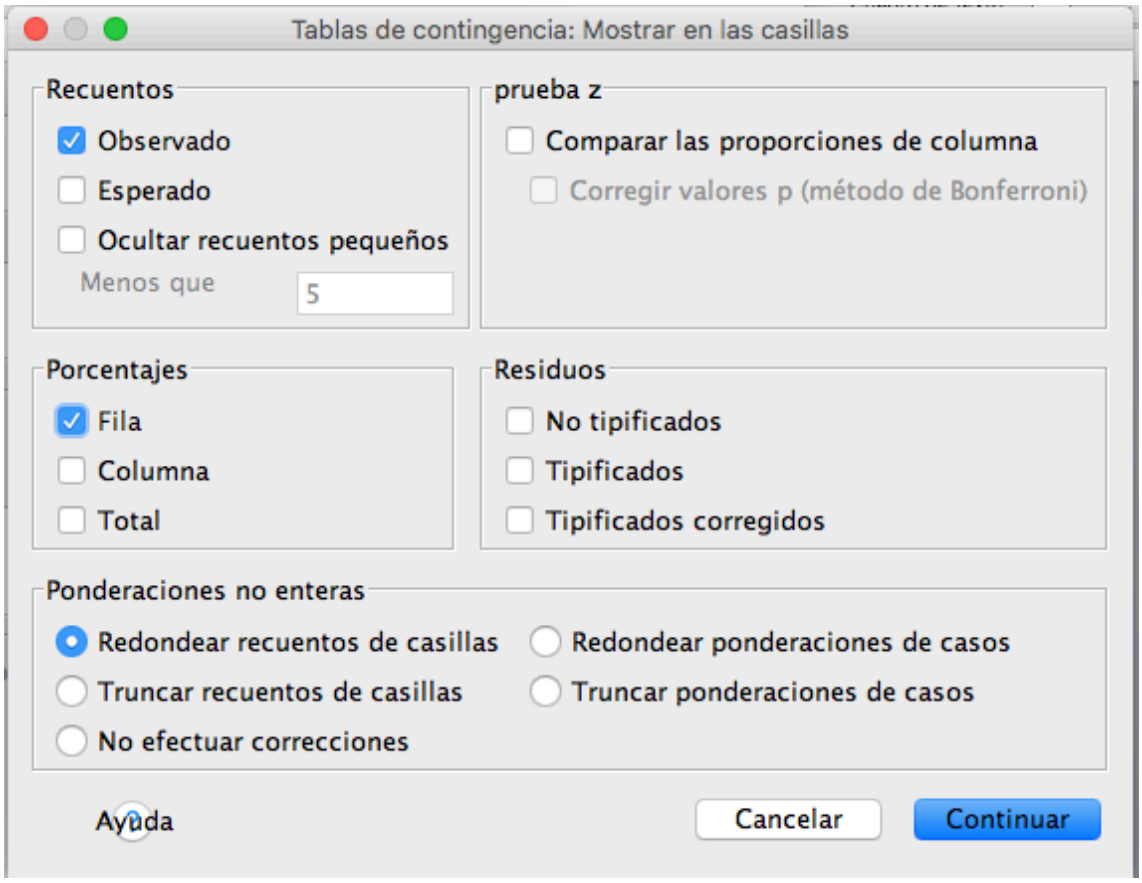
Filas: sexo (1=hombre, 6=mujer)
Columnas: alcohol (0=nunca, 1=bajo, 2=moderado/alto)

sexo * alcohol Crosstabulation

Count

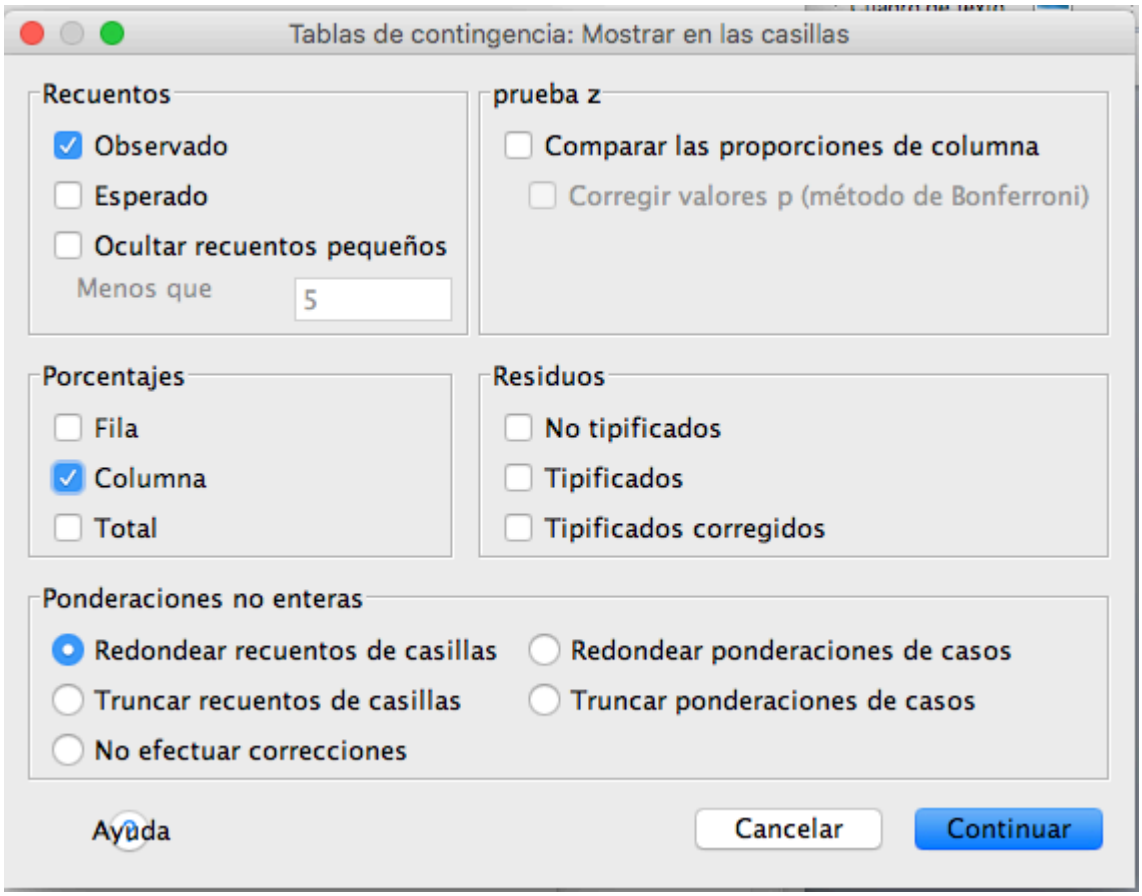
		alcohol			Total
		Nunca	Bajo	Moderado/Alto	
sexo	Hombre	8	36	48	92
	Mujer	16	79	13	108
Total		24	115	61	200

Se puede solicitar, pinchando en Casillas los porcentajes por filas o por columnas.



sexo * alcohol Crosstabulation

			alcohol			Total
			Nunca	Bajo	Moderado/Alto	
sexo	Hombre	Count	8	36	48	92
		% within sexo	8,7%	39,1%	52,2%	100,0%
	Mujer	Count	16	79	13	108
		% within sexo	14,8%	73,1%	12,0%	100,0%
Total		Count	24	115	61	200
		% within sexo	12,0%	57,5%	30,5%	100,0%

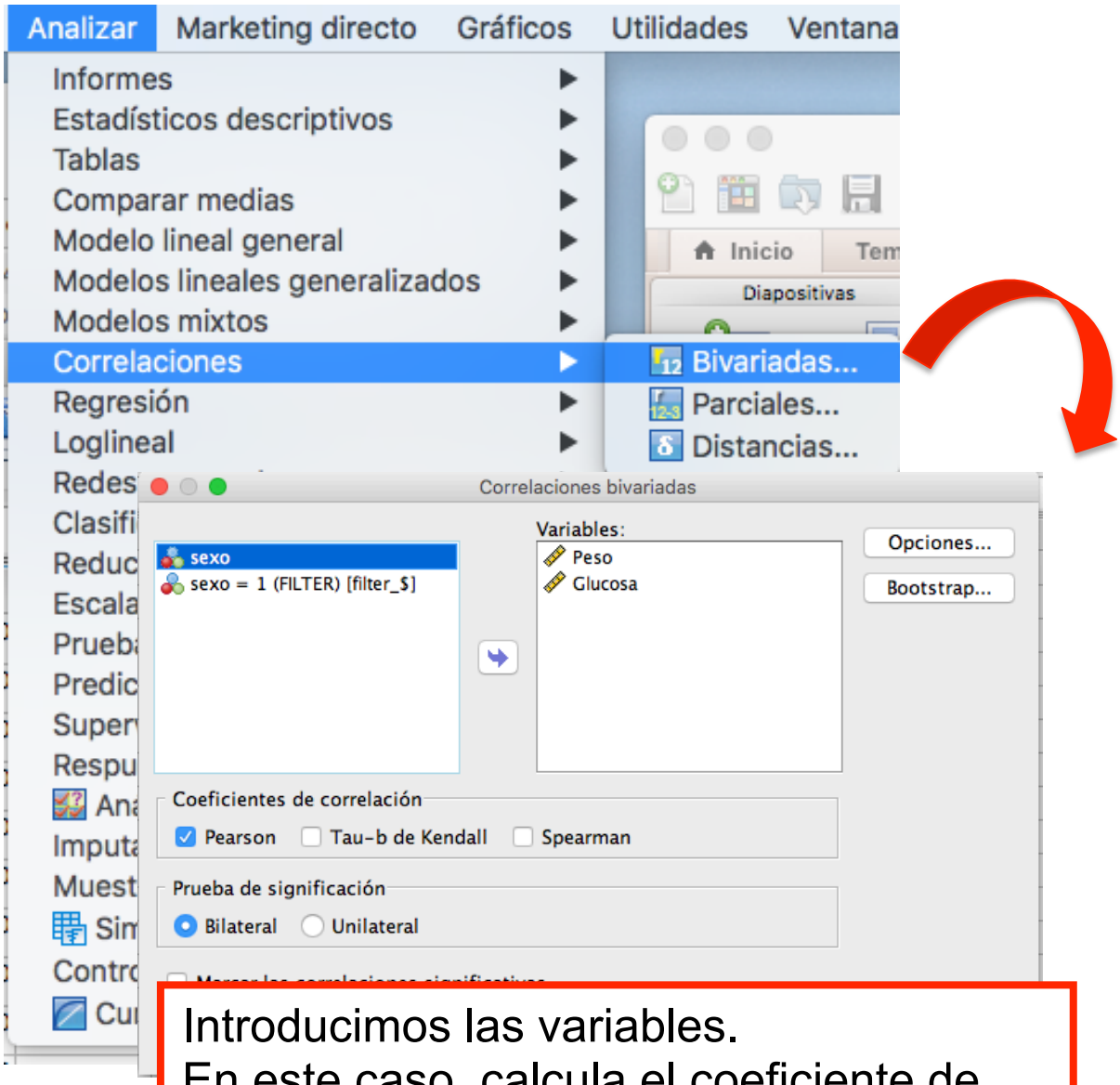


sexo * alcohol Crosstabulation

			alcohol			Total
			Nunca	Bajo	Moderado/Alto	
sexo	Hombre	Count	8	36	48	92
		% within alcohol	33,3%	31,3%	78,7%	46,0%
	Mujer	Count	16	79	13	108
		% within alcohol	66,7%	68,7%	21,3%	54,0%
Total		Count	24	115	61	200
		% within alcohol	100,0%	100,0%	100,0%	100,0%

6. Análisis de correlación

Coeficiente de correlación lineal

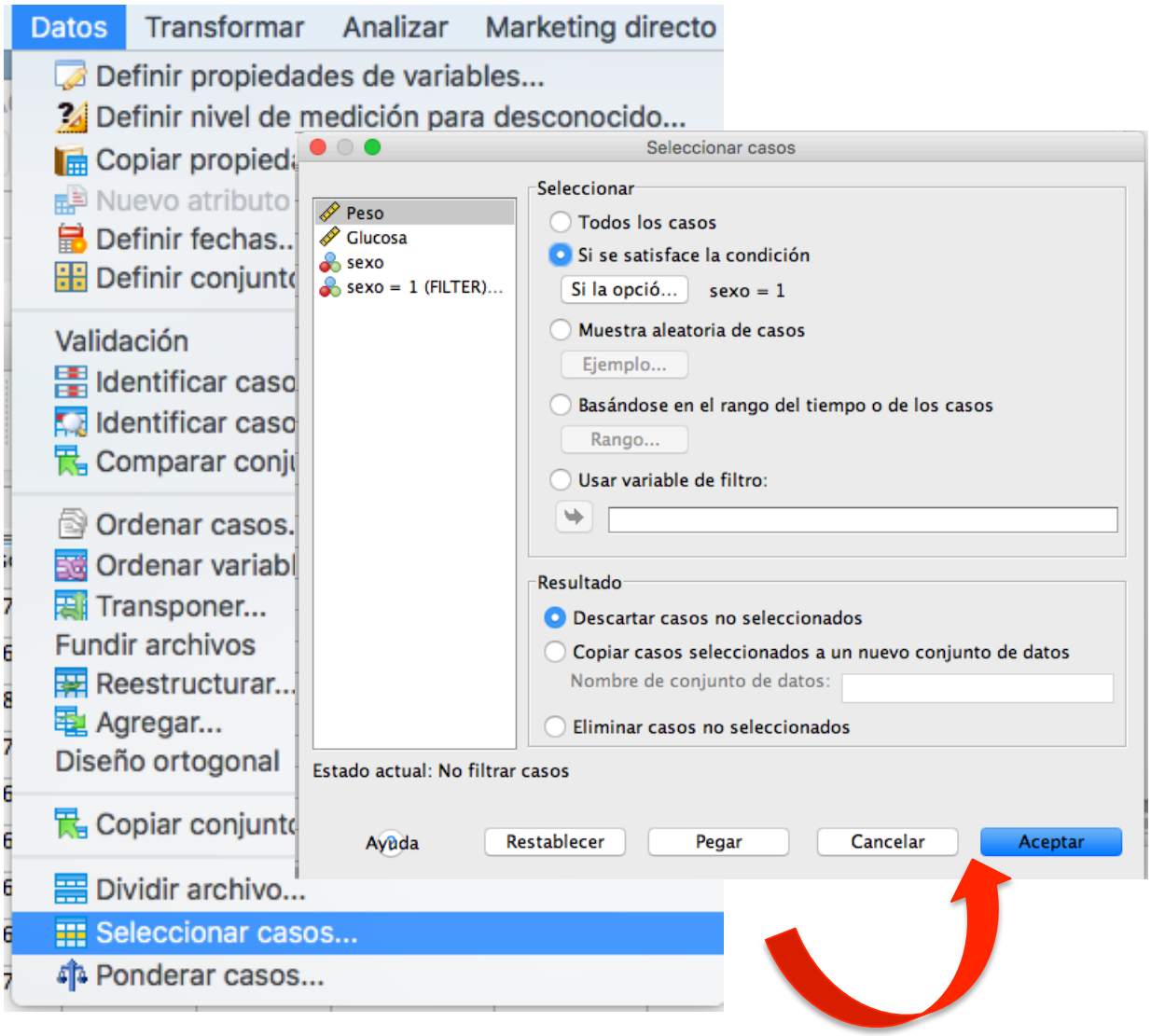


Introducimos las variables.
En este caso, calcula el coeficiente de correlación conjunto.

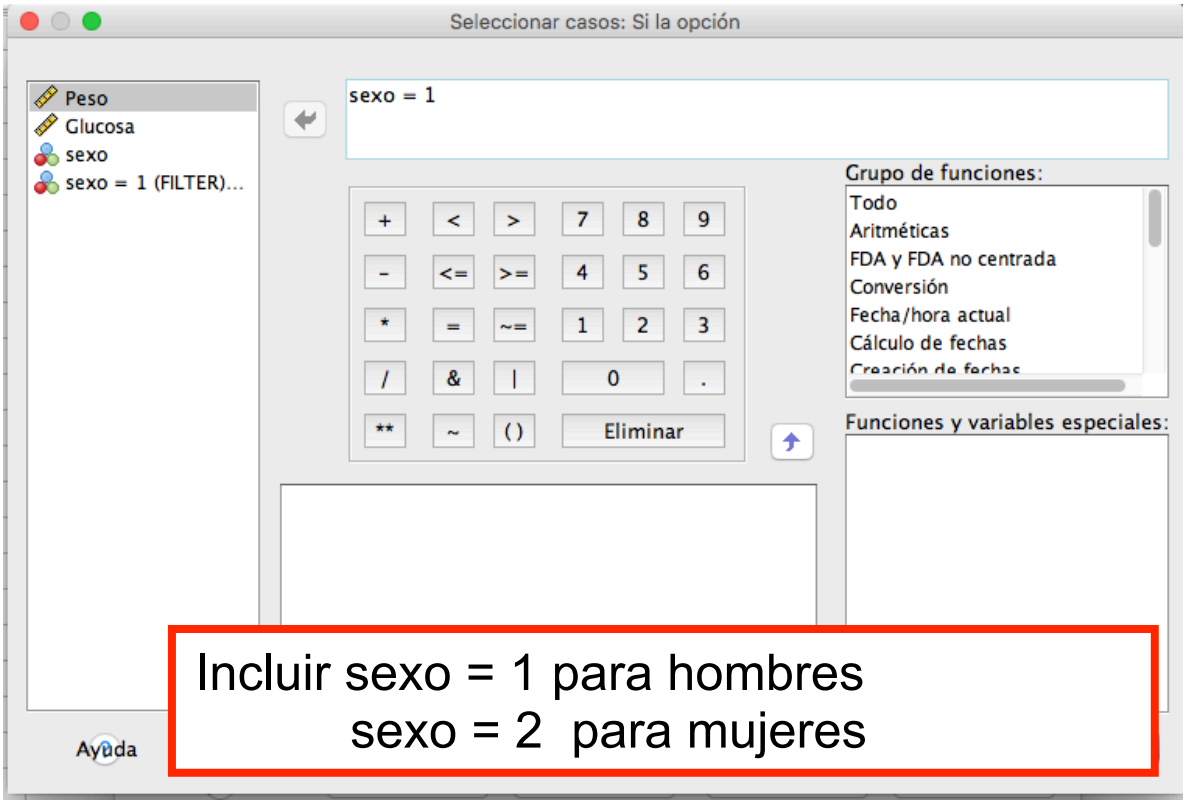
En el caso que quieras calcular por sexo, hay que seleccionar los casos o segmentar el archivo por sexo.

Correlations

		PESO KG	glucosa g/100ml
PESO KG	Pearson Correlation	1	,099
	Sig. (2-tailed)		,566
	N	36	36
glucosa g/100ml	Pearson Correlation	,099	1
	Sig. (2-tailed)	,566	
	N	36	36



Pulsar en “ Si se satisface la condición”



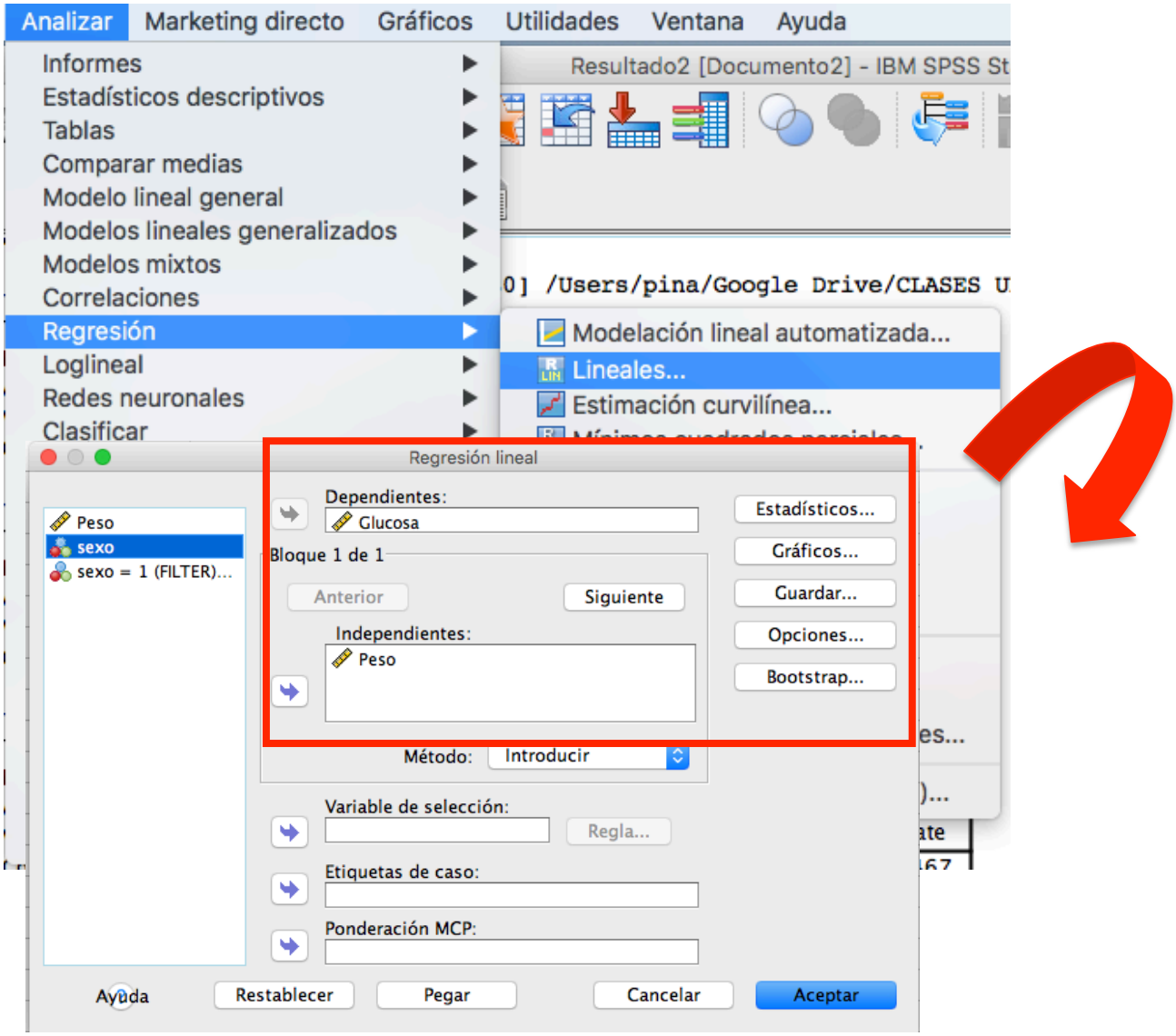
Correlations

		PESO KG	glucosa g/100ml
PESO KG	Pearson Correlation	1	,630**
	Sig. (2-tailed)		,005
	N	18	18
glucosa g/100ml	Pearson Correlation	,630**	1
	Sig. (2-tailed)	,005	
	N	18	18

**. Correlation is significant at the 0.01 level (2-tailed).

7. Análisis de regresión lineal

Dependiente: glucosa (variable continua)
Independiente: Peso(variable Continua)

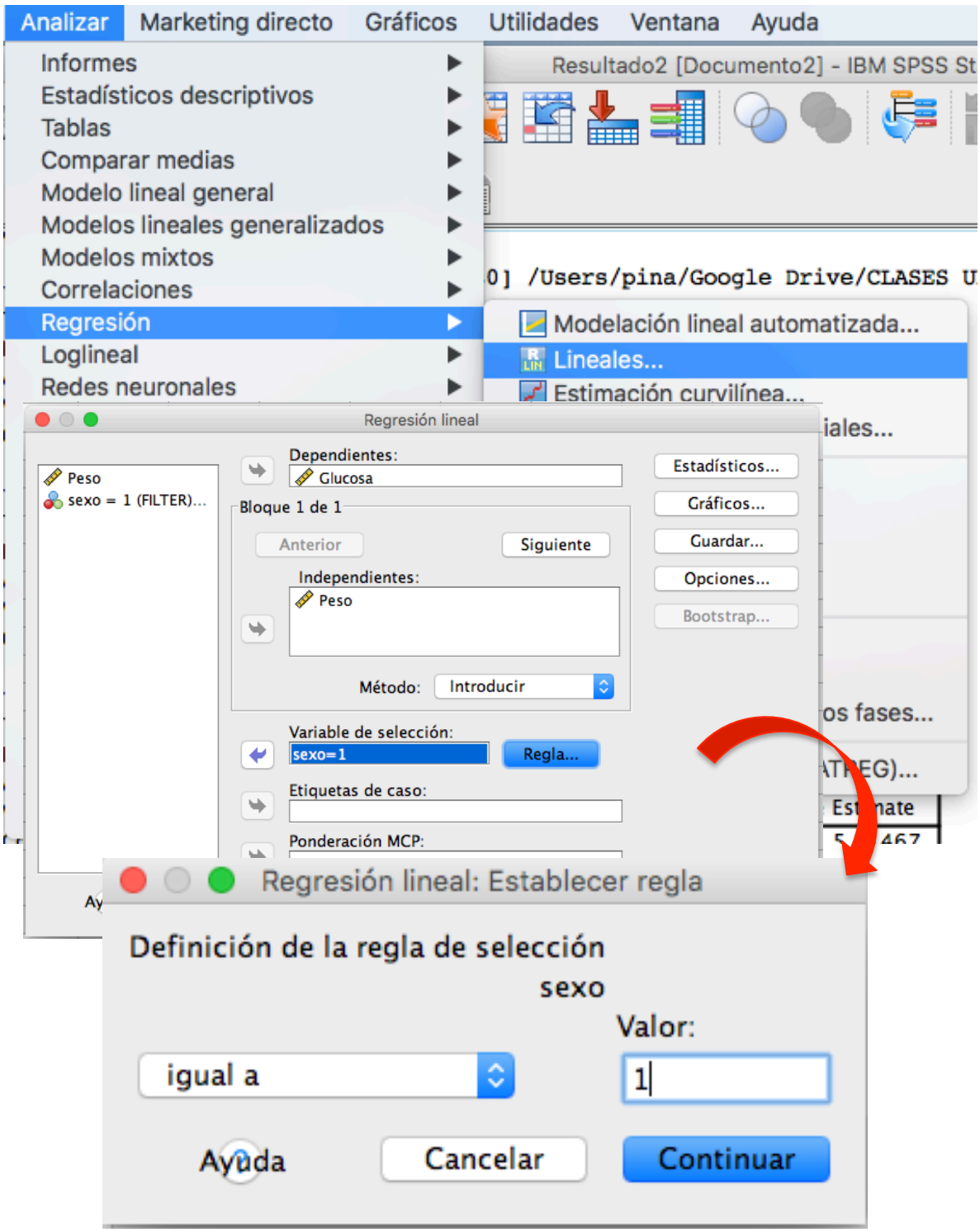


Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	96,719	9,770		9,900	,000
	PESO KG	,087	,150	,099	,580	,566

a. Dependent Variable: glucosa g/100ml

En el caso que desees calcular la regresión lineal sólo para hombres.



Coefficients[a,b]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	46,997	16,026		2,932	,010
	PESO KG	,734	,226	,630	3,247	,005

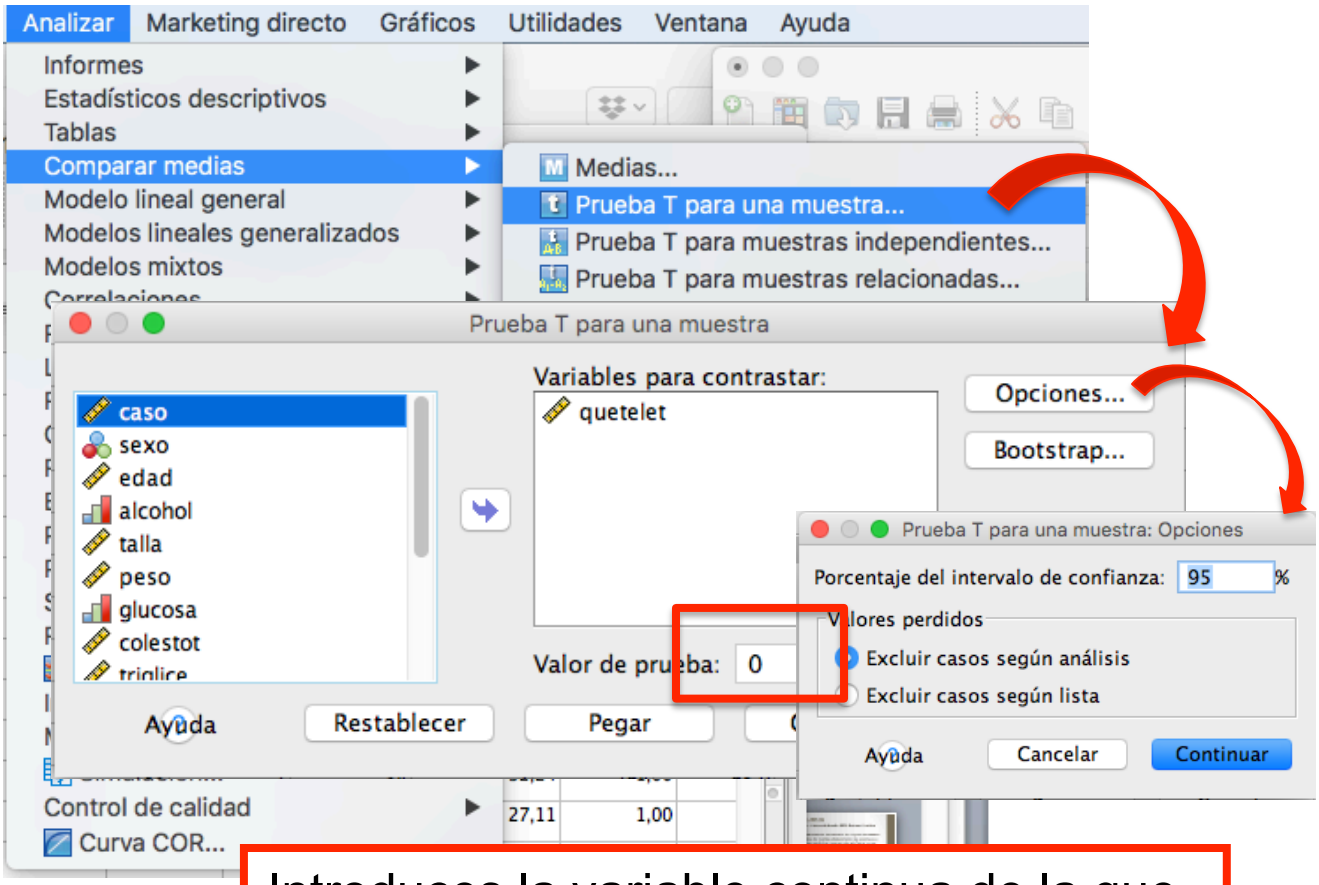
a. Dependent Variable: glucosa g/100ml

b. Selecting only cases for which SEXO PACIENTES = hombre

8. Intervalos de confianza

8.1 Intervalo de confianza para una media

Analizar medias → Prueba T para una muestra



Introduces la variable continua de la que deseas obtener el intervalo de confianza.

Después pulsas en **Opciones** para indicar el nivel de confianza.

En valor de prueba se deja "0" ya que el programa calcula un intervalo de confianza para $\mu-\mu_0$, es decir que $\mu_0=0$

Obtenemos la siguiente salida:

One-Sample Statistics

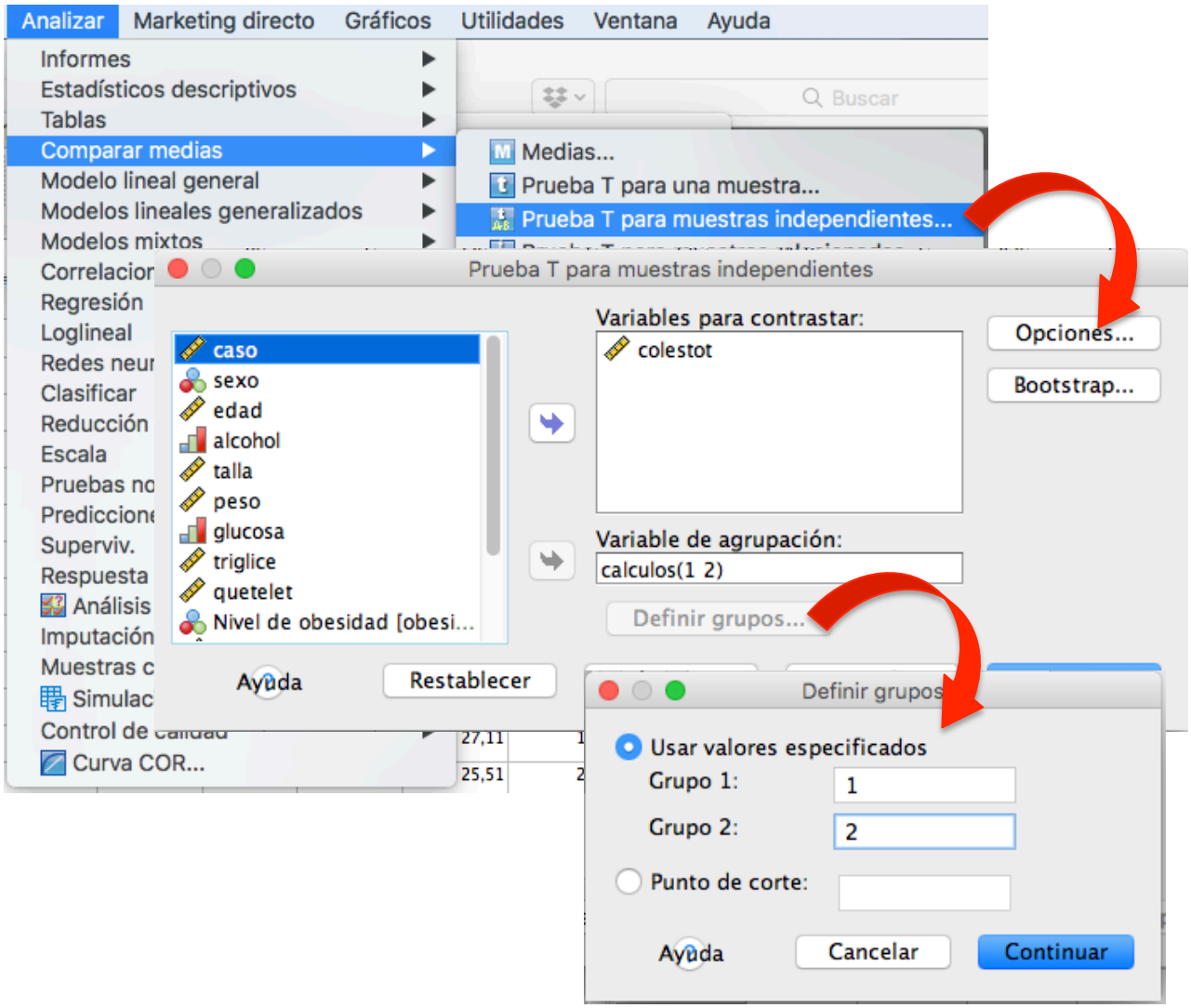
	N	Mean	Std. Deviation	Std. Error Mean
quetelet	200	26,7317	3,69368	,26118

One-Sample Test

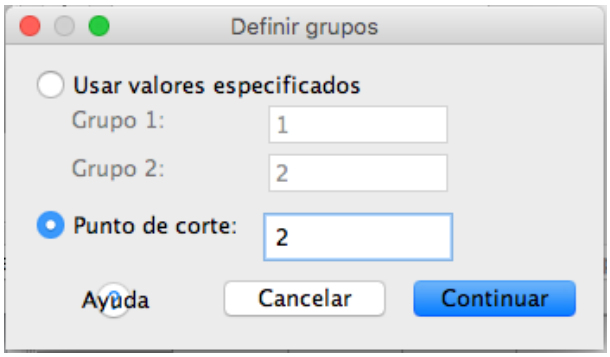
	Test Value = 0					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
quetelet	102,349	199	,000	26,73170	26,2167	27,2467

8.2 Intervalo de confianza para la diferencia de medias

Analizar medias → Prueba T para muestras independientes



En definir grupos, se indica el código asignado a cada grupo de la variable en cuestión. En opciones introduciremos el nivel de confianza.



En caso de que la variable de agrupación tenga 3 grupos, puede indicar un punto de corte y SPSS creará dos grupos:
>=2
<2

Obtenemos la siguientes salidas:

Group Statistics

	calculos	N	Mean	Std. Deviation	Std. Error Mean
colestot	No	183	225,51	51,184	3,784
	Si	17	247,53	40,033	9,709

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
colestot	Equal variances assumed	1,380	,242	-1,724	198	,086	-22,016	12,773	-47,203	3,172
	Equal variances not assumed			-2,113	21,185	,047	-22,016	10,421	-43,675	-,356

Group Statistics

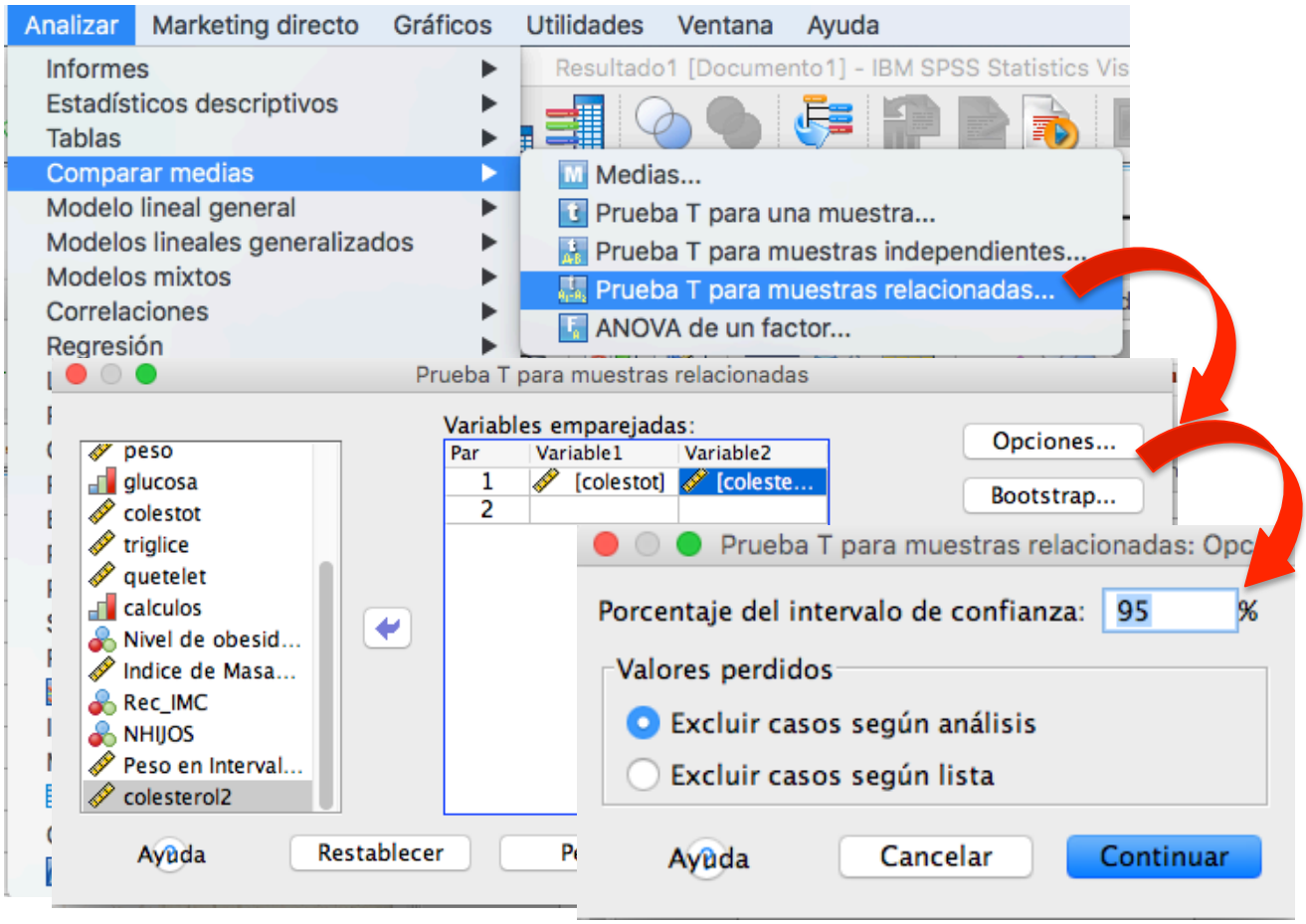
	Nivel de obesidad	N	Mean	Std. Deviation	Std. Error Mean
triglice	>= 2	131	141,12	99,553	8,698
	< 2	69	80,42	55,831	6,721

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
triglice	Equal variances assumed	12,498	,001	4,688	198	,000	60,702	12,949	35,167	86,237
	Equal variances not assumed			5,522	197,190	,000	60,702	10,992	39,024	82,379

8.3 Intervalo de confianza para la diferencia de medias en muestras relacionadas

Analizar medias → Prueba T para muestras relacionadas



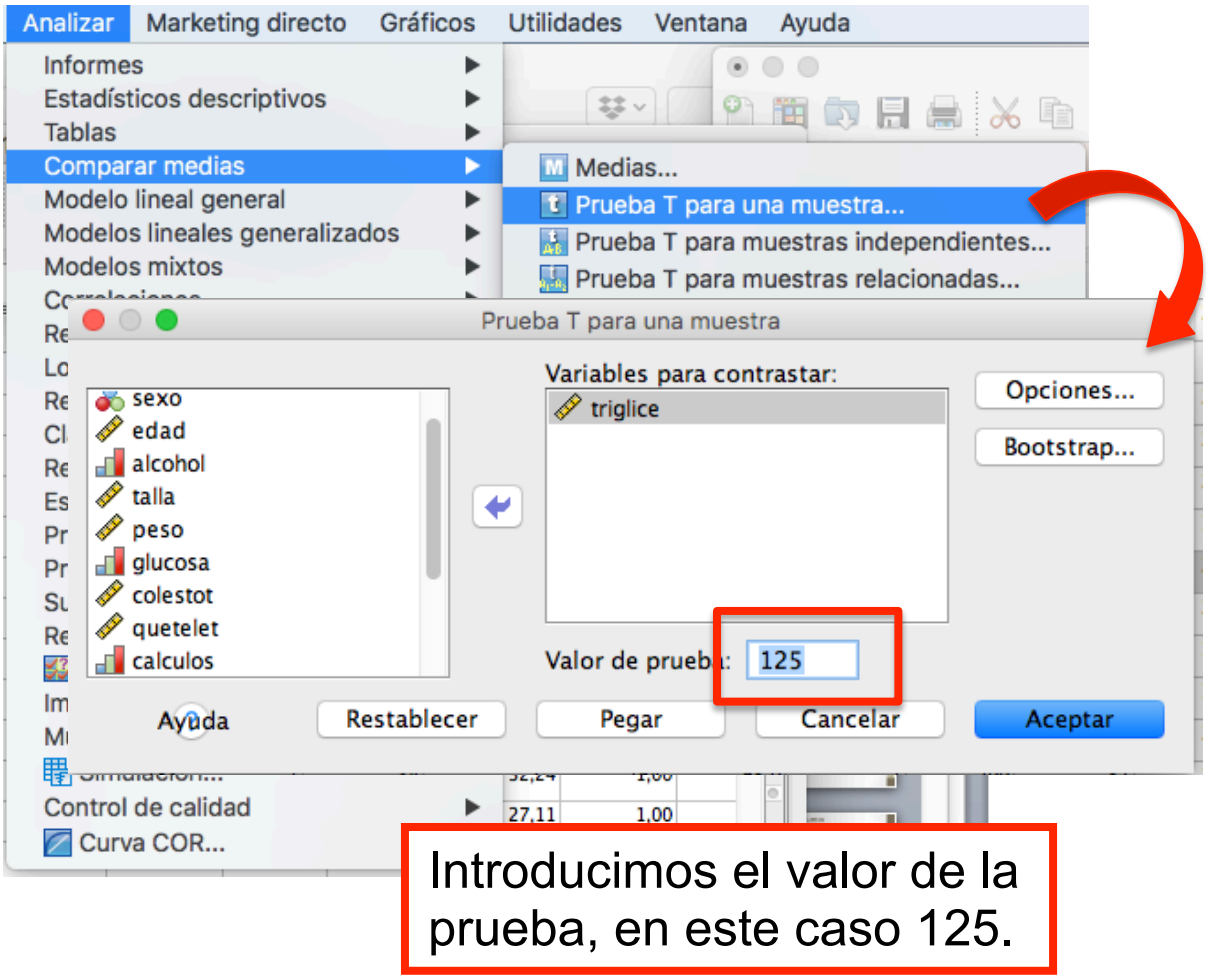
9. Contraste de hipótesis

9.1 Contraste de hipótesis para una media

Analizar medias → Prueba T para una muestra

$$H_0 : \mu = 125$$

$$H_a : \mu \neq 125$$



Obtenemos la siguiente salida:

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
triglice	200	120,18	91,523	6,472

One-Sample Test

	Test Value = 125					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
triglice	-,745	199	,457	-4,820	-17,58	7,94

Esta salida muestra los resultados del contraste de la *t de Student* con un intervalo de confianza para la diferencia entre el valor observado y el valor teórico (contrastado). Cada una de las columnas de la tabla muestra:

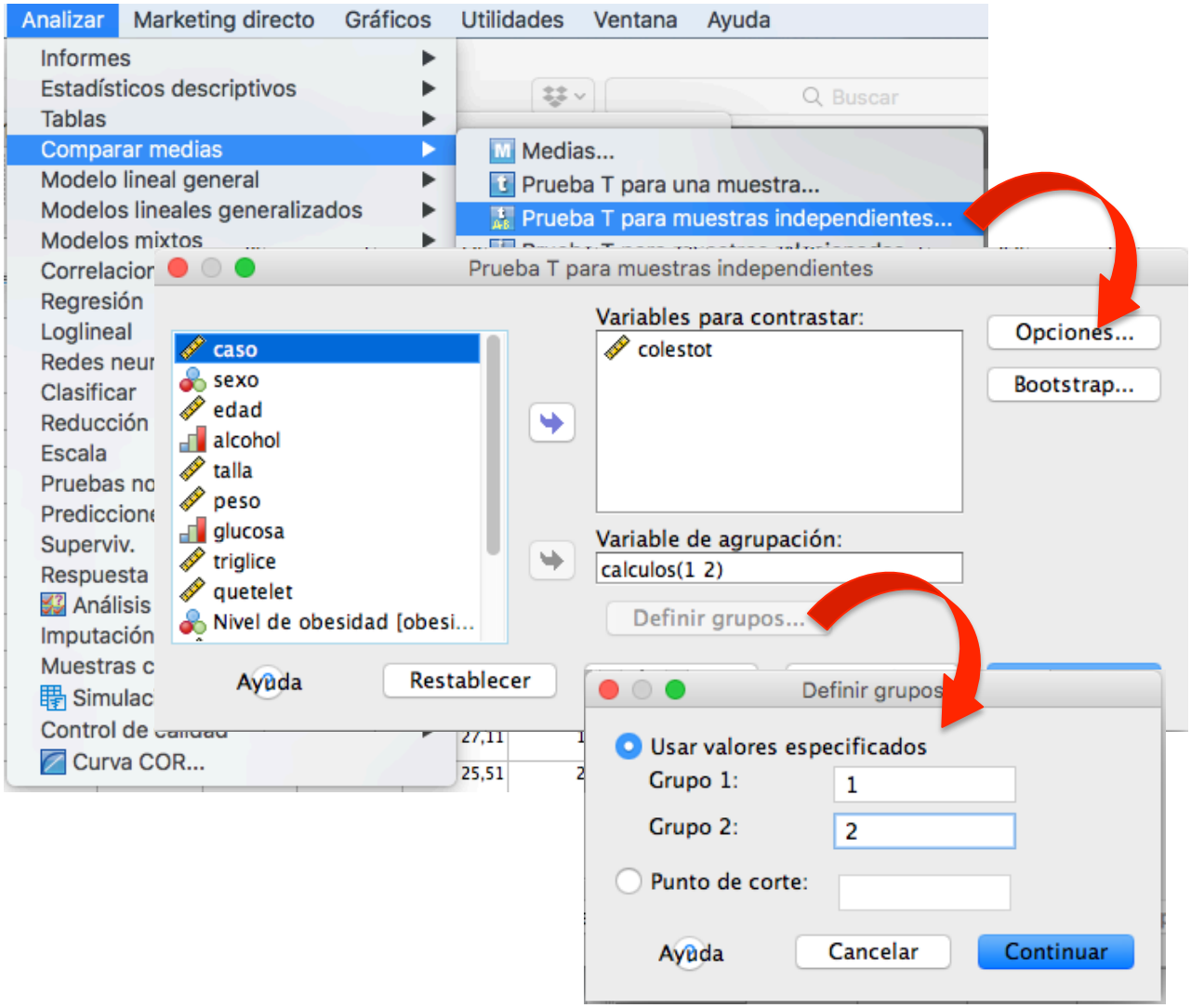
- t = -,745: El valor experimental del estadístico de contraste
- gl = 199: Los grados de libertad
- Sig.= 0,457: El p-valor o nivel crítico del contraste
- Diferencia de medias = -4.820 Es la diferencia entre la media teórica (125) y la media observada (120,18)
- 95% Intervalo de confianza = (-17,58 7,94): Es el intervalo de confianza para la diferencia entre la media teórica y la media observada al nivel de confianza del 95%.

9.2 Contraste de hipótesis para una diferencia de medias

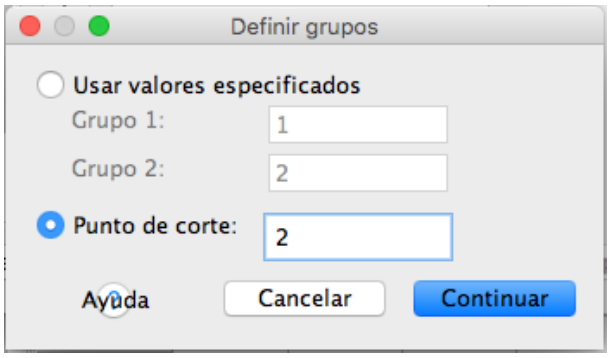
Analizar medias → Prueba T para muestras independientes.

$$H_0 : \mu - \mu_0 = 0$$

$$H_a : \mu - \mu_0 \neq 0$$



En definir grupos, se indica el código asignado a cada grupo de la variable en cuestión. En opciones introduciremos el nivel de confianza.



En caso de que la variable de agrupación tenga 3 grupos, puede indicar un punto de corte y SPSS creará dos grupos:
>=2
<2

Obtenemos la siguiente salida:

Group Statistics

	calculos	N	Mean	Std. Deviation	Std. Error Mean
colestot	No	183	225,51	51,184	3,784
	Si	17	247,53	40,033	9,709

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
colestot	Equal variances assumed	1,380	,242	-1,724	198	,086	-22,016	12,773	-47,203	3,172
	Equal variances not assumed			-2,113	21,185	,047	-22,016	10,421	-43,675	-,356

Para realizar un contraste de diferencia de medias de dos poblaciones independientes **hay que contrastar previamente las varianzas de dichas poblaciones**. Esta salida nos muestra el valor experimental del estadístico de contraste (Fexp = 1.280), este valor deja a la derecha un área igual a 0.242 (Sig.= 0. 242), por lo tanto no se puede rechazar la hipótesis nula de igualdad de varianzas.

A continuación se realiza el contraste para la diferencia de medias suponiendo que las varianzas son iguales. La tabla nos muestra el valor experimental del estadístico de contraste (texp = -1.724) y el p-valor = 0.086 (Sig.= 0.086), por lo tanto no se puede rechazar la hipótesis nula de igualdad de medias. También, se puede concluir el contraste observando que el intervalo de confianza para la diferencia de medias (-47,203, 3.172) contiene al cero.

9.3 Contraste de hipótesis Chi-cuadrado

La prueba de independencia Chi-cuadrado, nos permite **determinar si existe una relación entre dos variables categóricas**. Es necesario resaltar que esta prueba nos indica si existe o no una relación entre las variables, pero no indica el grado o el tipo de relación; es decir, no indica el porcentaje de influencia de una variable sobre la otra o la variable que causa la influencia.

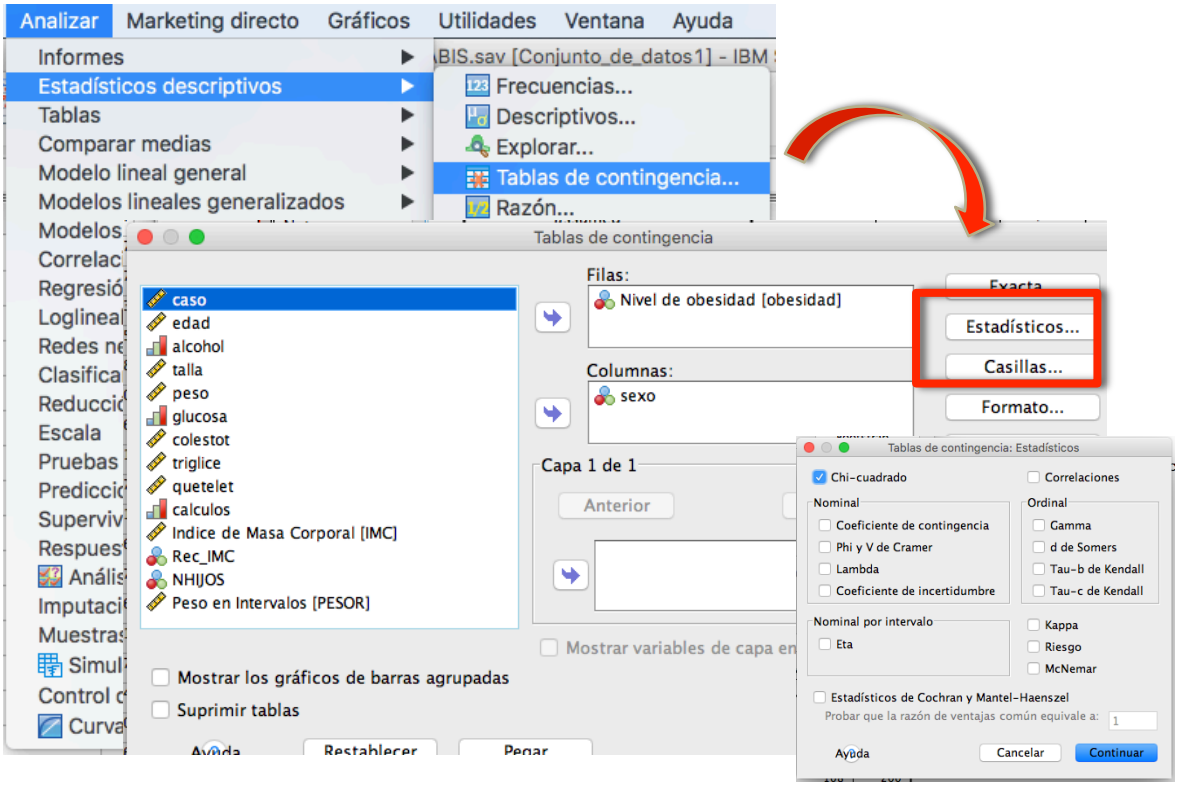
Analizar medias → Tablas de contingencia → Estadísticos

$$H_0 : \text{No existe asociación - proporciones iguales}$$

$$H_a : \text{Existe asociación - proporciones diferentes}$$

$$H_0 : \text{variables independientes}$$

$$H_a : \text{no son independientes}$$



Obtenemos la siguiente salida:

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Nivel de obesidad * sexo	200	100,0%	0	0,0%	200	100,0%

Nivel de obesidad * sexo Crosstabulation

Count

		sexo		Total
		Hombre	Mujer	
Nivel de obesidad	Delgado o normal (IQ<25)	32	37	69
	Sobrepeso (25<=IQ<30)	48	46	94
	Obesidad (IQ>=30)	12	25	37
Total		92	108	200

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	3,716[a]	2	,156
Likelihood Ratio	3,790	2	,150
Linear-by-Linear Association	1,107	1	,293
N of Valid Cases	200		

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 17,02.

La prueba de independencia del Chi-cuadrado, parte de la hipótesis que las variables son independientes; es decir, que no existe ninguna relación entre ellas y por lo tanto ninguna ejerce influencia sobre la otra. El objetivo de esta prueba es comprobar la hipótesis mediante el nivel de significación, por lo que sí el valor de la significación es mayor o igual que el *Alfa* (0.05), se acepta la hipótesis, pero si es menor se rechaza.

La tabla nos muestra el valor experimental del estadístico de contraste 3,716 y el p-valor = 0.156 (Sig.= 0.256), por lo tanto no se puede rechazar la hipótesis nula de independencia de variables, es decir que las dos variables son independientes.